THE

LONDON, EDINBURGH, AND DUBLIN

PHILOSOPHICAL MAGAZINE

AND

JOURNAL OF SCIENCE.

[SIXTH SERIES.]

JANUARY 1923.

I. *Dissociation-equilibria by the Method of Partitions.* By R. H. Fowler, *M.A., Fellow and Lecturer in Trinity College, Cambridge* *.

§ 1. *Introduction and Summary of the paper.*—In two previous papers † Mr. Darwin and I have treated the usual problems of the partition of energy in statistical mechanics by a more powerful method than has hitherto been used. This method allowed us to obtain all the usual formulæ of statistical equilibrium by the direct calculation of averages, avoiding entirely the need for any appeal to entropy. In the second paper we examined the whole position of entropy in statistical theory, and concluded that, though a proper logical account of it is possible, such an account must be somewhat recondite, and that the whole idea is really unnecessary in statistical work. Practically speaking, before entropy can be properly introduced the whole statistical problem for which it might be helpful has been otherwise solved.

These two papers apply only to assemblies of systems in which the various systems retain an individual existence throughout—that is to say, all cases of dissociation were

* Communicated by the Author.

† Phil. Mag. Sept. and Nov. 1922: "On the Partition of Energy," Parts I. and II. These two papers will be referred to as first and second paper respectively. As I need scarcely say, I am, in this paper, deeply indebted to Mr. Darwin, who must not, however, be held responsible for its imperfections.

specifically excluded. The method can, however, be extended, with some loss of elegance, to assemblies in which dissociation occurs. It is still, however, simple and satisfactory, and in addition leads to what I believe to be extensions of known results. The equilibrium of such assemblies treated by purely statistical methods forms the subject of this paper which is Part III. of the series.

The problems of dissociation have recently been treated by Ehrenfest and Trkal * from the statistical standpoint in a most interesting paper. They maintain the position that the study of chemical constants, thermodynamic probability, and absolute entropy, which is left in considerable obscurity by earlier writers †, can only be logically approached by a statistical study of the dissociation-equilibria of molecules and atoms. However, they only discuss the dissociation problem on semi-classical lines: they assume, that is, that all degrees of freedom are either "fully excited" so that the classical theory is adequate, or "completely unexcited" in which case they may be ignored. This restriction can be removed so as to treat the statistical problem of dissociation quite generally—generally that is to say from the mathematical point of view. The difficulties which are met with and restrict the range of the results are no longer mathematical but apparently physical and inherent in the problem.

In our second paper Mr. Darwin and I have already discussed in some detail the same questions of thermodynamic probability and absolute entropy, but of course without including cases of dissociation. With the position maintained by Ehrenfest and Trkal we are in complete agreement, and with the help of the results here obtained this discussion can be carried further.

It is convenient to outline here the scope of the present paper. In § 2 I summarize the results of our previous papers so far as they are required here. In §§ 3-6 I introduce the study of the dissociation problem by the discussion in detail of the simplest case, the reaction $X_2 \rightleftarrows 2X$. This suffices to bring out all the salient features of the method and the extensions to previous results which are possible by its aid. In § 7 the results are applied in

* Ehrenfest and Trkal, Proc. Sect. Sciences Amsterdam, vol. xxiii. p. 162 (1920).

† M. Planck, *Wärmestrahlung*, 3rd ed. §§ 127, *sqq.*; O. Sackur, *Ann. der Physik*, vol. xxxvi. p. 958 (1911), vol. xl. p. 67 (1913); H. Tetrode, *Ann. der Physik*, vol. xxxviii. p. 434 (1912); etc.

general terms to Saha's * theory of the ionization in stellar atmospheres. This is perhaps the most interesting application of the theory of dissociation-equilibria. Saha's fundamental formula is based on the whole apparatus of chemical constants and Nernst's Heat Theorem. In view of its importance, and the somewhat unsatisfactory nature of its basis, it is of interest to present this formula in its proper form as a theorem in statistical mechanics. In so doing, one can obtain a clearer insight into the exact assumptions on which it is based. One is led to the conclusion that the usual formula is not necessarily correct—a further term should be included whose effect may be of great importance. This extra term unfortunately may make Saha's formula less definite, and the detailed application of the new formula must be postponed for further investigation. The discussion of the temperature ionization of a gas introduces us to a fundamental difficulty in the statistics of the Hydrogen (or any other) atom †. This is considered in detail in § 8. In § 9 the results of §§ 3–6 are extended to the general dissociation problem.

In § 10 I consider by the same method the general vapour-pressure equation, which is also discussed by Ehrenfest and Trkal ‡. This leads in § 11 to general formulæ for the chemical constants, which show up their dependence on the weights of the lowest permitted quantum states of the molecules. In particular, it is now possible to discuss adequately the chemical constant of H_2 in the region in which it behaves effectively like a monatomic gas. The agreement of its chemical constant with both theory and experiment for genuine monatomic gases is shown to provide important additional evidence as to the weights to be assigned to the first two quantum states of the rotational motion. In particular, it confirms the absence of the state of no rotation demanded by the evidence of band-spectra in general and by the curve of its own specific heats. Finally, the constants in the dissociation equations are expressed in terms of the chemical constants introduced into the vapour-pressure equations. The whole of these relations can be presented, and are

* M. N. Saha, Phil. Mag. vol. xl. pp. 472, 809 (1920); Proc. Roy. Soc. vol. xcix. A. p. 135 (1921). For general expositions, see E. A. Milne, The Observatory, vol. xliv. p. 261 (Sept. 1921); H. N. Russell, Astroph. Journ. vol. lv. p. 119 (1922).

† K. F. Herzfeld, *Ann. der Physik*, vol. li. p. 262 (1916).

‡ *Loc. cit.*, additional notes No. III.

here presented, as theorems in pure statistical mechanics based on the general principles of the quantum theory.

It is natural to follow up these theorems with the extended discussion of entropy referred to above, and in particular of Nernst's heat theorem and its relationship to the weight of the lowest quantum state of the systems concerned. This has been already pointed out in a preliminary way*, but with the additional theoretical material now available the connexion can be brought out more clearly. This discussion must, however, form the subject of another paper—it is unduly long for inclusion here.

§ 2. *Summary of previous results on Partitions.*—In the previous papers by Mr. Darwin and myself we have expounded a method of evaluating the partition of energy for various types of assembly. The essential point of our method is the calculation of average values by means of integrals whose integrands are simply expressed in terms of "partition functions." There is a certain partition function for each type of system in the assembly—quantized or classical, the latter being obtained by a limiting process. By such means the tiresome combinatory expressions are very easily handled, and the integrals which give the average values and the fluctuations about those average values can be simply evaluated by the "method of steepest descents."

The partition function—which in a different notation is the "Zustandsumme" of Planck—is constructed as follows. The physically possible states of the system are enumerated. These form a (generally infinite) series of discrete states for a quantized system; for the systems of classical mechanics the available phase-space must be divided into cells, all of whose dimensions must ultimately be made to tend to zero. To each state attaches a weight. The weight of any cell of a system obeying Hamiltonian equations is proportional to its extension. The relative weights of the cells of a quantized system are (at any rate in simple cases) determinable by the use of Bohr's Correspondence Principle and Ehrenfest's Adiabatic Hypothesis. No further *a priori* principle appears to be required. All weights are then definitely fixed by the convention† that

* Second paper, § 8.

† This convention is the simplest and most convenient that can be made (first paper, § 2), though it differs from that of Ehrenfest and Trkal (*loc. cit.*) by a constant factor. But in no case can a constant factor in *all* the weights affect any physical result.

any one state of a simple quantized system shall have unit weight. It should be remembered, however *, that the only essential condition attaching to the weights is precisely analogous to that attaching to entropy in classical thermodynamics—a definite ratio is required (and must be assignable) between the weights of two states of a system (or systems) which can pass from one to the other, but as long as the two states (or systems) are mutually non-convertible their weights are non-comparable and may have any ratio whatever. In the dissociation problem a quantized system (a molecule) can turn into two or more classical systems (free atoms), and so the complete specification of relative weights outlined above is relevant and must be retained. It has been used for simplicity throughout these papers. It here, for the first time, plays an integral part in the formulæ.

Associated with any state or cell there is a definite energy depending on the cell and (possibly) on certain parameters defining the external fields of force. Then, if p_r is the weight and ϵ_r the energy in the rth state or cell for any particular system, the partition function is defined in terms of a certain additional parameter ϑ by the equation

$$f(\vartheta) = \Sigma_r p_r \vartheta^{\epsilon_r}, \quad . \quad . \quad . \quad . \quad . \quad (2{\cdot}1)$$

summed over all states or cells. It is a function of ϑ, x_1, x_2, ... For classical systems the sum must be replaced by its limit when all the dimensions of all the cells tend to zero, that is by an integral ; but this limit can be used in all formulæ without restriction †, exactly as if it were still of the form (2·1).

We recall certain examples of partition functions.

(i.) Planck's line vibrator :

$$f(\vartheta) = \frac{1}{1-\vartheta^{\epsilon}}, \quad (\epsilon = h\nu). \quad . \quad . \quad . \quad . \quad . \quad (2{\cdot}21)$$

(ii.) The rotations of a diatomic molecule of transverse moment of inertia J ‡ :

$$f(\vartheta) = 1 + 3\vartheta^{\epsilon} + 5\vartheta^{4\epsilon} + 7\vartheta^{9\epsilon} + \ldots, \quad \left(\epsilon = \frac{h^2}{8\pi^2 \mathrm{J}}\right). \quad (2{\cdot}22)$$

* Second paper, § 8.

† First paper, § 12. For further details, see a forthcoming paper in Proc. Camb. Phil. Soc. (1922).

‡ The weight $2n+1$ assigned to the nth quantum state must on general principles be correct if none of the apparently possible states are barred. This was pointed out in our first paper, § 11. It is so to speak the maximum value of the weight. But there is a good

(iii.) Free monatomic molecules of mass m in a volume V:

$$f(\vartheta) = \frac{(2\pi m)^{3/2}\mathrm{V}}{h^3(\log 1/\vartheta)^{3/2}}. \quad . \quad . \quad . \quad . \quad . \quad (2{\cdot}23)$$

If any system is capable of two or more types of motion which are independent of one another—for example, the motions of translation and rotation of a di- or polyatomic molecule—the partition functions for the two types can be calculated separately and their product is the partition function for the whole motion of the system.

With the help of these partition functions we construct an integral for C, the total number of *weighted complexions**, which enumerates all possible states of any assembly of systems whose total energy is E. If there are $\mathrm{M_A}$ systems of type A with partition function f_A etc., then

$$\mathrm{C} = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z^{\mathrm{E}+1}}\left[f_\mathrm{A}(z)\right]^{\mathrm{M_A}}\left[f_\mathrm{B}(z)\right]^{\mathrm{M_B}}\ldots\ldots; \quad . \quad (2{\cdot}3)$$

the contour γ is any circle round the origin of radius less than unity. Similar integrals can at once be obtained for $\mathrm{C}\overline{\mathrm{E}}_\mathrm{A}$, where $\overline{\mathrm{E}}_\mathrm{A}$ is the average energy held by the systems A, or $\mathrm{C}\bar{a}_r$, where $\bar{a}_r$ is the average number of systems A in their rth quantum state.

These integrals are easily evaluated approximately (E and the M's all large). It is found that there is a *unique* real number ϑ, $0<\vartheta<1$, determined by the equation

$$\mathrm{E} = \Sigma_\mathrm{A}\mathrm{M_A}\vartheta\frac{d}{d\vartheta}\log f_\mathrm{A}, \quad . \quad . \quad . \quad . \quad (2{\cdot}31)$$

and that if the contour γ passes through ϑ the whole value of the integral is contributed by a small part of the contour near $z=\vartheta$. We thus find

$$\overline{\mathrm{E}}_\mathrm{A} = \mathrm{M_A}\vartheta\frac{d}{d\vartheta}\log f_\mathrm{A}, \quad . \quad . \quad . \quad . \quad (2{\cdot}32)$$

$$\bar{a}_r = \mathrm{M_A}p_{\mathrm{A},r}\,\vartheta^{\epsilon_{\mathrm{A},r}}/f_\mathrm{A}. \quad . \quad . \quad . \quad . \quad (2{\cdot}33)$$

---

deal of evidence, some of which we must consider in later sections, that not all these states can be permissible. The first term may be and probably is missing, and the later weights may be $2n$ instead of $2n+1$. See Reiche, *Ann. der Physik*, vol. lviii. p. 657 (1919). Reiche considers also n and $n+1$ as possible values; but, as we shall see in § 11, these cannot be allowed if full agreement in the limit with classical mechanics is to be maintained. Whatever the precise form of $f(\vartheta)$ its treatment here is unaffected.

* We use complexions in an extended sense to cover the case in which we have already proceeded to the limit for classical systems. For the legitimacy of this, see reference in last footnote.

We then identify ϑ with the temperature on an empirical scale and prove that in all cases it must be related to the absolute temperature T by the equation

$$\vartheta = \epsilon^{-1/kT}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (2{\cdot}4)$$

where k is Boltzmann's constant.

The average values above are shown to be normal properties of the assembly by evaluating their fluctuations $\overline{(E_A - \overline{E_A})^2}$ etc., and showing that $\overline{(E_A - \overline{E_A})^2}$ is of the same order as $\overline{E_A}$, therefore negligible compared with $(\overline{E_A})^2$.

§ 3. *The Dissociation-equilibrium of the reaction* $X_2 \rightleftarrows 2X$.

To simplify the formulæ and exhibit in the clearest way the important points in the argument, I have thought it best to begin by discussing the simplest case, in which only one kind of atom is present and only one reaction can occur. In any such problem, as Ehrenfest and Trkal point out, it is not really allowable to treat the internal vibrations of the diatomic molecule along its axis of symmetry as completely unexcited*. This degree of freedom must be allowed its proper share of energy, which may be appreciable if there is appreciable dissociation. We must also of course include the kinetic energy of translation of both atoms and molecules and the rotations of the latter about any transverse axis. Axial rotations of the molecules and all rotations of the atoms can be ignored as completely unexcited. There is no difficulty in formally including them if required.

Let us suppose that the partition function for the combined transverse rotations and vibrations of the molecule X_2 is, in general terms, $b(z)$ †. It is not properly possible to divide $b(z)$ into factors for the rotations and vibrations separately, for the two motions are not independent. We shall discuss shortly the details of the specification of $b(z)$ in § 4. Further, let $h_1(z)$ be the partition function for the kinetic energy of translation of the atoms (mass m) and $h_2(z)$ for the molecules (mass $2m$) ; $h_1(z)$ and $h_2(z)$ are given by (2·23).

The number of weighted complexions c for an assembly of M atoms X and N molecules X_2 in a volume V with energy F available for distribution is by (2·3),

$$c = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z^{F+1}}[h_1(z)]^M[h_2(z)b(z)]^N. \quad . \quad . \quad (3{\cdot}1)$$

* It is by including this and similar quantized motions that the results of these authors are extended and completed in this paper.

† Ultimately $b(z)$ is determined by and determines the band spectrum of the molecule.

This, however, is no longer the total number of complexions of which the assembly is capable ; for M and N can change. We must take account of all possible values of M and N and all possible ways in which the assembly with given M and N can be constructed out of a total of X atoms. This part of the problem is completely discussed by Ehrenfest and Trkal*. They show that the total number of essentially different ways in which an assembly of M atoms and N molecules can be built up out of a total of X atoms is

$$\frac{X!}{M!\,N!\,\sigma^N}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (3{\cdot}2)$$

where

$$X = M+2N, \quad . \quad . \quad . \quad . \quad . \quad (3{\cdot}21)$$

and σ is the symmetry number † of the molecule, which in this case is 2. We shall retain the symbol σ in the calculations.

Since there are $X!/\{M!\,N!\,\sigma^N\}$ examples of the assembly with M atoms and N molecules each represented by c weighted complexions, the full number of complexions representing the assembly in this statistical state is

$$\frac{X!\,c}{M!\,N!\,\sigma^N}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (3{\cdot}3)$$

The total C of all complexions representing the assembly is obtained by summing (3·3) over all values of M and N subject to (3·21) and the condition of constant total energy. The energy F available for distribution among the partition functions when there are M atoms and N molecules depends on M and N. If χ is the heat of dissociation per molecule, or the work required to separate one molecule into two independent atoms, we can express F in terms of χ. For the sake of preciseness, let us take the state of zero energy of the assembly as that in which the molecules are all dissociated

* *Loc. cit.* §4, equation (18). It is in the treatment of what Ehrenfest and Trkal call the "μ-weight" that our discussions diverge, allowing of increased generality here.

† The symmetry number of a molecule (Ehrenfest and Trkal, *loc. cit.* p. 163) is the number of different orientations which it can take up without altering the distribution in space of any type of atom. The expression $X!/(M!\,N!)$ gives the number of arrangements of the X atoms to form M atoms and N molecules, and arrangements in which the molecules X_1X_2 and X_2X_1 occur are counted as different. But owing to the rotational freedom of the molecule these arrangements are not really distinct, and the above expression must therefore be divided by 2^N. For a molecule with tetrahedral symmetry such as CH_4, $\sigma=12$, and so on.

and the resulting X atoms all at rest. The assembly can then have negative energies, but this is of no importance. Let us take the state of zero energy of a molecule to be its state when it has no rotation * or velocity of translation and no internal energy of vibration. It is then in its lowest quantum state, and χ must be taken to be the work required to separate such a molecule into two independent atoms at rest. In the ordinary process in which atoms combine to form molecules with evolution of heat, $\chi > 0$. Observe that χ is an absolute constant. The fact that not all molecules are in their lowest quantum state does not alter χ. This fact is taken full account of by the partition function $b(z)$. If E is the total energy of the assembly (relative to our conventional zero), then

$$\mathrm{F} = \mathrm{E} + \mathrm{N}\chi, \quad . \quad . \quad . \quad . \quad . \quad . \quad (3{\cdot}4)$$

and therefore

$$\mathrm{C} = \Sigma_{\mathrm{M}+2\mathrm{N}=\mathrm{X}} \frac{\mathrm{X}!}{\mathrm{M}!\,\mathrm{N}!\,\sigma^{\mathrm{N}}} \frac{1}{2\pi i} \int_\gamma \frac{dz}{z^{\mathrm{E}+1}} [h_1(z)]^{\mathrm{M}} \left[\frac{h_2(z)\,b(z)}{z^\chi}\right]^{\mathrm{N}}$$

$$= \frac{1}{2\pi i} \int_\gamma \frac{dz}{z^{\mathrm{E}+1}} [h_1(z)]^{\mathrm{X}} \Sigma_{\mathrm{N}} \frac{\mathrm{X}!\,\beta^{\mathrm{N}}}{\mathrm{N}!\,(\mathrm{X}-2\mathrm{N})!}, \quad . \quad . \quad . \quad (3{\cdot}5)$$

where

$$\beta = \frac{h_2(z)\,b(z)}{\sigma z^\chi \{h_1(z)\}^2}. \quad . \quad . \quad . \quad . \quad . \quad (3{\cdot}51)$$

Exactly similar integrals can be obtained at once for $\mathrm{C}\overline{\mathrm{N}}$ etc., which give us the degree of dissociation and its fluctuations if required. Thus by (3·3),

$$\mathrm{C}\overline{\mathrm{N}} = \frac{1}{2\pi i} \int_\gamma \frac{dz}{z^{\mathrm{E}+1}} [h_1(z)]^{\mathrm{X}} \Sigma_{\mathrm{N}} \frac{\mathrm{N}\mathrm{X}!\,\beta^{\mathrm{N}}}{\mathrm{N}!\,(\mathrm{X}-2\mathrm{N})!}. \quad (3{\cdot}52)$$

We show how to evaluate these and similar integrals in § 5. We give in advance the principal result, which is that

$$\frac{\overline{\mathrm{N}}}{(\mathrm{X}-2\overline{\mathrm{N}})^2} = \frac{\overline{\mathrm{N}}}{(\overline{\mathrm{M}})^2} = \beta(\vartheta) = \frac{h_2(\vartheta)\,b(\vartheta)}{\sigma\vartheta^\chi \{h_1(\vartheta)\}^2}. \quad . \quad (3{\cdot}6)$$

Equation (3·6) gives the dissociation-equilibrium, "the reaction-isobar," as a function of the temperature. Its physical form is discussed shortly in the next section.

* The arguments are easily adapted to include the case in which this state is not permitted; see § 4.

§ 4. *The form of the Reaction-isobar for* $X_2 \rightleftarrows 2X$.— If we insert known values for the functions $h(\vartheta)$ and for σ in (3·6) we get

$$\frac{V\bar{N}}{(\bar{M})^2} = \frac{h^3}{2(\pi m)^{3/2}} \frac{(\log 1/\vartheta)^{3/2}}{\vartheta\chi} b(\vartheta)$$

$$= \frac{h^3}{2(\pi km)^{3/2}} T^{-3/2} e^{\chi/kT} B(T). \quad . \quad . \quad . \quad (4·1)$$

Expressed in terms of molecular densities $\nu_1 = \bar{M}/V$, $\nu_2 = \bar{N}/V$, we have

$$\frac{\nu_2}{\nu_1^2} = \frac{h^3}{2(\pi km)^{3/2}} T^{-3/2} e^{\chi/kT} B(T). \quad . \quad . \quad . \quad (4·11)$$

As we have assumed that the individual gases (of molecules and of atoms) are perfect, we can transform to partial pressures p_1, p_2 by $p_1 = \nu_1 kT$, $p_2 = \nu_2 kT$. If, further, x is the fraction of molecules dissociated and P the total pressure, then

$$p_1 = \frac{2xP}{1+x}, \quad p_2 = \frac{(1-x)P}{1+x},$$

$$\frac{1-x^2}{4x^2P} = \frac{h^3}{2(\pi m)^{3/2} k^{5/2}} T^{-5/2} e^{\chi/kT} B(T). \quad . \quad (4·12)$$

The essentially new feature in these equations is the occurrence in its proper place of the factor $b(\vartheta)$ or B(T).

The term $b(\vartheta)$ is the partition function for the rotational and vibrational energy of the molecules. As we have already said, these cannot strictly be separated. But as a first approximation these motions are independent and the separation can be made, as is shown by the evidence of band spectra. Thus

$$b(\vartheta) = r(\vartheta)\, v(\vartheta).$$

We suppose that J the moment of inertia is roughly constant to the order of accuracy with which we can discuss $b(\vartheta)$; then $r(\vartheta)$ is given by (2·22) unless the rotationless molecule is impossible *, in which case the term 1 must be omitted. When $\vartheta \to 0$ $(T \to 0)$,

$$r(\vartheta) \to 1, \quad r(\vartheta) \backsim p_1 e^{-\epsilon/kT}, \quad (p_1 = 2 \text{ or } 3), \quad . \quad (4·2)$$

* There is strong evidence from band spectra (Sommerfeld, *Atombau und Spectrallinien*, ed. 3, chap. 7) and specific heats (Reiche, *loc. cit.*) that this state is impossible.

according as the term 1 is or is not to be retained. When $\vartheta \to 1$ ($T \to \infty$), in either case *

$$r(\vartheta) \to \frac{8\pi^2 J k T}{h^2}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4{\cdot}21)$$

giving the usual classical contribution R to the molecular specific heats.

For the vibrational term $v(\vartheta)$ one may suppose for a similar rough approximation that the vibrations are like those of Planck's line vibrator, but that as an energy χ will dissociate the molecule no more than p states are possible in which the molecule remains a molecule. Then

$$v(\vartheta) = 1 + \vartheta^{v} + \ldots + \vartheta^{(p-1)v} = \frac{1-\vartheta^{pv}}{1-\vartheta^{v}}. \quad . \quad . \quad (4{\cdot}3)$$

It is somewhat less arbitrary to suppose that we may take

$$v(\vartheta) = 1 + \vartheta^{v_1} + \ldots + \vartheta^{v_{p-1}}, \quad (0 < v_1 < v_2 < \ldots < v_{p-1} < \chi); \quad . \quad . \quad . \quad (4{\cdot}31)$$

the weights are necessarily all unity. When $\vartheta \to 0$ ($T \to 0$), $v(\vartheta) \to 1$. When $\vartheta \to 1$ ($T \to \infty$), $v(\vartheta) \to p$.

With reference to large values of T, it should, however, be observed that if χ and therefore pv are fairly large compared with v, there will be a considerable interval of values of ϑ or T in which ϑ^{v} is nearly 1 while ϑ^{pv} is still small, and in fact negligible compared with 1. Thus, for such values of T that † $pv \gg T \gg v$,

$$v(\vartheta) = \frac{1}{1-\vartheta^{v}} = \frac{kT}{v}, \quad \frac{\vartheta v'(\vartheta)}{v(\vartheta)} = \frac{v}{\vartheta^{-v}-1} = kT,$$

approximately. We can of course write $v = h\nu_0$, where ν_0 is the fundamental frequency of the band spectrum. Thus at such temperatures the specific heat at constant volume of our molecular gas may reach effectively the value $\frac{7}{2}R$ or 6·96 calories per gram-molecule ‡ in spite of the limitation to p states. A similar result must hold for the $v(\vartheta)$ given by (4·31) if the v_p do not differ much from multiples of v.

* By a theorem of Cesaro's: Bromwich, 'Infinite Series,' p. 132. Equation (4·21) holds for the weights $2n$ or $2n+1$, but not of course for the weight n.

† The notation $x \gg y$ is coming into increasing use in Mathematical Physics to denote "x is large compared with y." It is a convenient and significant symbolism, which is adopted in this paper.

‡ For example, C_v for the molecule I_2 is stated to reach this value at about 600° C. Stern, *Ann. der Physik*, vol. xliv. p. 512 (1914).

The evidence of band spectra shows that in general the υ-terms are considerably greater than the rotation ϵ-terms $\left(\epsilon=\frac{h^2}{8\pi^2 \mathrm{J}}\right)$; so that the rotations can be regarded as fully excited, while the vibrations are still practically unexcited. In such a case the formula for the reaction-isobar takes successively the approximate forms

$$\frac{1-x^2}{4x^2\mathrm{P}} = \frac{h^3}{2(\pi m)^{3/2}k^{5/2}}\mathrm{T}^{-5/2}e^{\chi/k\mathrm{T}}, \quad (k\mathrm{T}\ll\epsilon), \quad . \quad . \quad . \quad (4{\cdot}41)$$

or

$$\frac{ph^3}{2(\pi m)^{3/2}k^{5/2}}\mathrm{T}^{-5/2}e^{(\chi-\epsilon)/k\mathrm{T}}, \quad (p_1=2 \text{ or } 3), \quad . \quad (4{\cdot}42)$$

$$= \frac{4\pi^{1/2}h\mathrm{J}}{(km)^{3/2}}\mathrm{T}^{-3/2}e^{\chi/k\mathrm{T}}, \quad (\epsilon\ll k\mathrm{T}\ll\chi,\ \upsilon), \quad . \quad . \quad (4{\cdot}43)$$

$$= \frac{4\pi^{1/2}\mathrm{J}}{\nu_0 k^{1/2}m^{3/2}}\mathrm{T}^{-1/2}e^{\chi/k\mathrm{T}}, \quad (\upsilon\ll k\mathrm{T}\ll\chi), \quad . \quad . \quad (4{\cdot}44)$$

$$= \frac{4p\pi^{1/2}h\mathrm{J}}{(km)^{3/2}}\mathrm{T}^{-3/2}e^{\chi/k\mathrm{T}}, \quad (\chi\ll k\mathrm{T}). \quad . \quad . \quad . \quad . \quad (4{\cdot}45)$$

Equation (4·42) refers to the case in which no rotation is impossible; the molecule of lowest quantum grade has no longer no energy, but energy ϵ. Since χ retains its old meaning, $\chi-\epsilon$ is now the heat of dissociation per molecule. Thus (4·41) and (4·42) essentially differ only in the constant factor p_1, which arises from the different weights which must be assigned in these cases to the lowest quantum state. Equation (4·43) corresponds to the conditions considered by Ehrenfest and Trkal, and agrees with their result, but only if the weight of the nth quantum rotation is effectively $2n$ (or $2n+1$) for large values of n. Weights such as n or $n+1$, which can only be maintained by very forced arguments, make (4·43) differ from Ehrenfest and Trkal by a factor 2, and are therefore almost certainly wrong.

The complete constant factor in these equations plays a part which is necessarily equivalent to the part played by the chemical constants of the various atoms and molecules. The chemical constants, therefore, must essentially depend on the weights of the lowest possible quantum states ((4·41) and (4·42)). The importance of the weights in this connexion appears to be inadequately appreciated, for the precise assumptions underlying the usual treatment are somewhat difficult to trace. It is better to postpone discussion of this point to § 11 after the equation of vapour-pressure has been considered at the close of this paper.

We have here displayed the various obvious approximate forms which we may expect the reaction-isobar to assume. Equation (4·12) is presumably, however, quite general, and the factor B(T) should be very largely determinable from a study of the band spectrum—in particular, the band absorption spectrum. If this could be completely analysed at any one temperature, all the possible rotational and vibrational states of the molecule would be determined. It would then only remain to attach precise values to the weights to complete the specification of B(T). It is just possible that a combined study of the reaction-isobar and the band spectrum might prove of considerable interest.

In conclusion, a possible objection may be raised to all the reasoning on which these results are based—namely, that it ignores radiation, whereas in fact a change of quantum state in the molecule probably seldom or never occurs without the emission or absorption of the appropriate radiation. It is, however, perfectly possible to include the temperature radiation in the statistical discussion, as has been shown elsewhere *. The results as functions of the temperature are unaffected, and the omission of an explicit account of the radiation in equilibrium with the dissociating gas is thus justified. The equilibrium is the same (as it must be) whether exchanges of energy take place by radiation or by collisions alone ; but of course this does not imply that the equilibrium remains unaltered when the assembly is subjected to radiation of a different temperature from outside.

§ 5. *Mathematical discussion of the Dissociation Formulæ.*— The evaluation of the integrals of § 3 proceeds just as in the simple cases of our first paper, with the addition that the asymptotic form of the series under the integral sign must be determined. If we write the integrand in the integral (3·5) for C in the form

$$\Sigma_N \frac{X!}{N!\,M!\,\sigma^N} [h_2(z)b(z)]^N [h_1(z)]^M z^{-N\chi-E},$$

we see at once that every separate term is a product of power series † which obey all the conditions laid down in § 6 of our first paper. All that is there asserted of such power series can therefore equally be asserted of the sum Σ. Therefore the integrand on any circle γ has its maximum

* Darwin & Fowler, Proc. Camb. Phil. Soc. (1922).

† We ignore throughout the fact that $h_1(z)$, $h_2(z)$ are finally to be the limits of such series. The necessary refinements are easily made.

modulus on the real axis. and on the real axis has a unique minimum between 0 and 1 which is the sole such root of the equation

$$\mathrm{E} = \mathrm{X}\vartheta \frac{d}{d\vartheta} \log h_1(\vartheta) + \vartheta \frac{d}{d\vartheta} \log \Sigma_{\mathrm{N}} \frac{\mathrm{X}!\,\beta^{\mathrm{N}}}{\mathrm{N}!\,(\mathrm{X}-2\mathrm{N})!}. \quad (5{\cdot}1)$$

We can then evaluate C to the approximation required when E and X are large by making the circle γ pass through $z=\vartheta$, and integrating over a small part of the circle in this neighbourhood. Further progress, however, obviously depends on a knowledge of the asymptotic form of $\log \Sigma_{\mathrm{N}}$.

This series and all similar series that occur in such problems can be evaluated for real β by a study of the series in the neighbourhood of its maximum term. Moreover, only real values of β need be so studied ; for the function represented by the series is an analytic function of the complex variables β or z near the positive real axis. To evaluate the integrals for C etc., we require the approximate values of this function on the circle γ near its intersection with the positive real axis. For this purpose it is sufficient to know the values of the first two (and sometimes more) differential coefficients of the function along γ at this point. These, however, are known when we know the corresponding differential coefficients along the real axis for real values of β or z.

For any real value of z there is a unique maximum among the terms of the series ; for the ratio of consecutive terms of the series starts by being approximately $\beta\mathrm{X}^2$ and steadily diminishes to β/X. If we refer to (4·1), we see that $\beta = f(z)/\mathrm{V}$, where $f(z)$ is a function of z alone and independent of the number of systems in the assembly. Now we are considering the state of large assemblies, for which X is large. We have, therefore, to suppose that V is large, so that the ratio X/V is fixed. But in spite of this we can certainly arrange for X to be so large that $\beta\mathrm{X}^2 > 1$ and $\beta/\mathrm{X} < 1$ for any given value of z. Thus the terms start by increasing, and ultimately diminish without ever being able to increase again. There therefore exists a maximum which is unique.

The maximum occurs where two consecutive terms are as nearly equal as possible—that is, where

$$\mathrm{N} = \beta(\mathrm{X}-2\mathrm{N})(\mathrm{X}-2\mathrm{N}-1), \quad . \quad . \quad . \quad (5{\cdot}11)$$

or with sufficient accuracy

$$\mathrm{N}/\mathrm{M}^2 = \beta(z). \quad . \quad . \quad . \quad . \quad . \quad (5{\cdot}12)$$

In the further discussion the use of Stirling's formula appears to be for the first time unavoidable, but its use is easy to justify rigorously. We write

$$N = N_0 + \alpha, \quad M = M_0 - 2\alpha,$$

where α is an integer, and find that near the maximum term N_0, M_0, the terms of the series can be expressed in the form

$$\frac{1}{(2\pi)^{\frac{1}{2}}}\left(\frac{X}{M_0 N_0}\right)^{\frac{1}{2}} \exp\left\{-\tfrac{1}{2}\alpha^2\left(\frac{1}{N_0}+\frac{4}{M_0}\right)+N_0\log\beta+X\log X-X\right.$$
$$\left.-N_0\log N_0+N_0-M_0\log M_0+M_0\right\}. \quad . \quad (5{\cdot}2)$$

All other terms in (5·2) are of negligible order so long as α is confined to a certain stretch of values on either side of zero. This stretch can, however, be chosen so as to allow $\alpha\left(\frac{1}{N}+\frac{4}{M}\right)^{\frac{1}{2}}$ to range effectively from $-\infty$ to $+\infty$. The terms in this strip provide the whole effective contribution to the series, and can be replaced by the integral of (5·2) with respect to α taken from $-\infty$ to $+\infty$. If we then write down $\log \Sigma_N$ and omit small terms, we obtain the asymptotic equality (X, V large)

$$\log \Sigma_N \frac{X!\,\beta^N}{N!\,(X-2N)!} = N_0\log\beta+X\log X-X$$
$$-N_0\log N_0+N_0-M_0\log M_0+M_0. \quad (5{\cdot}3)$$

The integral (3·52) for CN can be treated in the same way. We find at once, to the same accuracy,

$$\Sigma_N \frac{NX!\,\beta^N}{N!\,(X-2N)!} = \beta\frac{d}{d\beta}\Sigma_N\frac{X!\,\beta^N}{N!\,(X-2N)!}$$
$$= N_0\Sigma_N\frac{X!\,\beta^N}{N!\,(X-2N)!}. \quad . \quad . \quad (5{\cdot}4)$$

If, now, we differentiate (5·3) with respect to β, we obtain

$$\beta\frac{d}{d\beta}\log\Sigma_N = N_0+\beta\frac{\partial N_0}{\partial\beta}(\log\beta-\log N_0)-\beta\frac{\partial M_0}{\partial\beta}\log M_0$$
$$= N_0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5{\cdot}41)$$

on using (5·12) and (3·21), which agrees with (5·4). This

proves that the asymptotic relation (5·3) can be differentiated with respect to β, and therefore also with respect to z.

Returning now to the integrals for C and $C\overline{N}$, it is easy to see how they are to be evaluated with the help of these asymptotic formulæ, and in particular that

$$\overline{N} = N_0, \quad . \quad . \quad . \quad . \quad . \quad . \quad (5{\cdot}5)$$

as determined by (5·12), provided that β is evaluated for the value of ϑ determined by (5·1). Thus the result $N/(\overline{M})^2 = \beta(\vartheta)$ is obtained, which we have already discussed in § 4. We can now express (5·1) itself in more intelligible form, for it becomes, using (5·3) and (5·41),

$$E = X\vartheta \frac{d}{d\vartheta} \log h_1(\vartheta) + \overline{N}\vartheta \frac{d}{d\vartheta} \log \beta(\vartheta), \quad . \quad . \quad (5{\cdot}6)$$

or

$$E = \overline{M}\vartheta \frac{d}{d\vartheta} \log h_1(\vartheta) + \overline{N}\vartheta \frac{d}{d\vartheta} \{\log h_2(\vartheta) + \log b(\vartheta)\} - \overline{N}\chi. \quad . \quad . \quad . \quad (5{\cdot}61)$$

This is precisely the form which the corresponding equation would take for an assembly in which N and M were *fixed* at their average values. We prove in the next section that the same is true of all details of the partition. The identification of ϑ with $e^{-1/kT}$ can be made by any of the methods of our former papers.

The changes in these results when the energy of radiation is accounted for may be mentioned here. It is found that the radiation has a partition function $R(z)$ which in the limit takes the form

$$R(z) = \exp\left\{\frac{8\pi^5 V}{45c^3h^3(\log 1/z)^3}\right\}, \quad . \quad . \quad . \quad (5{\cdot}7)$$

where c is the velocity of light. The only alteration in the formulæ is that the integrals for C etc. all contain a factor $R(z)$ in the integrand, and the equation (5·61) becomes

$$E = \overline{M}\vartheta \frac{d}{d\vartheta} \log h_1(\vartheta) + \overline{N}\vartheta \frac{d}{d\vartheta} \{\log h_2(\vartheta) + \log b(\vartheta)\} + \vartheta \frac{d}{d\vartheta} \log R(\vartheta) - N\chi. \quad . \quad (5{\cdot}71)$$

As functions of ϑ, the dissociation and the details of the energy partition are unaffected.

§ 6. *Further details of the partition.*—We have still to show that the possession of the values $\bar{\mathrm{N}}$ and $\bar{\mathrm{M}}$ is a "normal" property of the assembly; for this we should evaluate $\overline{(\mathrm{N}-\bar{\mathrm{N}})^2}$. We have at once

$$\mathrm{C}\overline{\mathrm{N}^2} = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z^{\mathrm{E}+1}}[h_1(z)]^{\mathrm{X}}\left(\beta\frac{d}{d\beta}\right)^2 \Sigma_{\mathrm{N}}\frac{\mathrm{X}!\,\beta^{\mathrm{N}}}{\mathrm{N}!\,(\mathrm{X}-2\mathrm{N})!}, \quad (6{\cdot}1)$$

and therefore, if we retain only terms of the highest order in X,

$$\overline{\mathrm{N}^2} = (\bar{\mathrm{N}})^2.$$

It follows at once that $\overline{(\mathrm{N}-\bar{\mathrm{N}})^2}$, which is equal to $\overline{\mathrm{N}^2}-(\bar{\mathrm{N}})^2$, is of lower order than $(\bar{\mathrm{N}})^2$, since the highest order terms cancel. This is sufficient by itself to prove that the possession of $\bar{\mathrm{N}}$ is a normal property of the assembly. The actual evaluation of $\overline{(\mathrm{N}-\bar{\mathrm{N}})^2}$, though straightforward, is somewhat tedious, and we shall not attempt it here *.

We can discuss in the usual way the partition of energy. For example, consider the average amount of energy taken up by the internal motions and rotations of the molecules $[b(z)]$. In any particular example of the assembly with M atoms and N molecules the sum of all possible values of the energy on the $b(z)$ terms is

$$\frac{1}{2\pi i}\int_\gamma \frac{dz}{z^{\mathrm{E}+1}}[h_1(z)]^{\mathrm{M}}[h_2(z)]^{\mathrm{N}}\left\{z\frac{d}{dz}[b(z)]^{\mathrm{N}}\right\}z^{-\mathrm{N}\chi}.$$

This follows at once by comparison with our first paper; for the sum required is the same as $c\bar{e}_b$, say, where $\bar{e}_b$ is the mean value of E_b for this special example of the assembly. Summing for all examples of the assembly, we get

$$\mathrm{C}\overline{\mathrm{E}_b} = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z^{\mathrm{E}+1}}[h_1(z)]^{\mathrm{X}}\left\{z\frac{d}{dz}\log b(z)\right\}\Sigma_{\mathrm{N}}\frac{\mathrm{N}\mathrm{X}!\,\beta^{\mathrm{N}}}{\mathrm{N}!\,(\mathrm{X}-2\mathrm{N})!}, \quad . \quad . \quad . \quad (6{\cdot}2)$$

and therefore

$$\overline{\mathrm{E}_b} = \bar{\mathrm{N}}\vartheta\frac{d}{d\vartheta}\log b(\vartheta), \quad . \quad . \quad . \quad . \quad (6{\cdot}21)$$

where ϑ is given by (5·61). Similar results hold for the other energy terms. We thus find, as above, that the partition of energy is the same as if N and M were fixed at their average values.

* For all such questions see Darwin and Fowler, Proc. Camb. Phil. Soc. (Mich. 1922).

Consider further the distribution of molecules among the various states included in $b(z)$. Let p_r and ϵ_r be the weight and energy corresponding to the rth quantum state, and a_r the number of molecules in that state. In a particular example of the assembly the sum of all possible values of a_r is*

$$\frac{1}{2\pi i}\int_\gamma \frac{dz}{z^{E+1}}\{Np_r z^{\epsilon_r}\}[h_1(z)]^M[h_2(z)]^N[b(z)]^{N-1}z^{-NX}.$$

Summing for all examples of the assembly, we get

$$C\overline{a_r} = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z^{E+1}}[h_1(z)]^X\left\{\frac{p_r z^{\epsilon_r}}{b(z)}\right\}\Sigma_N \frac{NX!\,\beta^N}{N!\,(X-2N)!}. \quad (6{\cdot}3)$$

Therefore

$$\overline{a_r} = \overline{N}p_r\vartheta^{\epsilon_r}/b(\vartheta), \quad . \quad . \quad . \quad . \quad . \quad (6{\cdot}31)$$

with the usual value of ϑ. Results similar in every way hold for the $h(z)$ factors †.

It is not necessary to suppose that the atoms and molecules of the X group are the only constituents of the assembly. We can suppose that any other systems in any quantity are also present, and, provided that no other reaction can occur, the extension of the whole of the preceding work is immediate. We can, for example, suppose that a constant volume perfect gas thermometer is also part of the assembly ; if we do this we have at once one means of identifying ϑ with $e^{-1/kT}$.

It is interesting, in conclusion, to view the whole equilibrium condition from a somewhat different standpoint. For simplicity, we will think of the rotational and vibrational terms as separated, so that the rotational terms have their own separate partition function. Then, instead of thinking of the molecules as all of one kind with p different internal energies, we can regard them as p different types of molecule, between each two of which types there is a proper statistical equilibrium. This equilibrium by (6·31) takes the form

$$\overline{N}_1/\overline{N}_2 = (p_1/p_2)e^{-\chi_{12}/kT}, \quad . \quad . \quad . \quad . \quad (6{\cdot}4)$$

where χ_{12} is the excess of internal energy in the state 1 over that in state 2. This is here the reaction-isobar parallel to (4·1) for the case of dissociation. Moreover, it is easy to see that (6·4) is precisely what (4·1) must reduce to in a

* First paper, §§ 8, 9.
† Second paper, § 12.

reaction in which the total number of individuals does not change. The equilibrium conditions between the different types of molecules are exactly parallel to reaction-equilibria, and may be treated as such.

§ 7. *The temperature ionization of Hydrogen and similar problems.*—The degree of ionization in a gas of H-atoms * in statistical equilibrium is precisely analogous to the dissociation problem already treated †. It is convenient to discuss it here on account of its interest as the basis of Saha's work on stellar atmospheres ; it raises also a practical difficulty of importance. It is convenient to consider the H-atom alone for the moment, since the structure and possible states of the H-atom may be regarded as completely known.

Consider an assembly of X H-nuclei and X electrons. These are now the elementary constituents, and the neutral H-atom is a complex—the molecule of our previous discussion—which can possess internal energy according to the orbit in which its electron is rotating. The masses of H and H^+ are the same, M, and the mass of the electron m. There are no relevant atomic rotations, and the symmetry numbers are all 1. If $b(z)$ is the partition function for the internal energy, we find, just as before, that the equilibrium state is

$$\bar{N}/(\bar{M})^2 = \beta(\vartheta), \quad \beta(\vartheta) = \frac{b(\vartheta)}{\vartheta^{\chi} h(\vartheta)}, \quad \vartheta = e^{-1/kT}, \qquad (7{\cdot}1)$$

where N is the number of neutral atoms, M the number of free electrons or H-nuclei, χ the ionization potential in ergs, and $h(\vartheta)$ the partition function for the motion of an electron.

When we come to specify $b(\vartheta)$ we meet at once a fundamental difficulty. With the conventions of § 3 for zero energy, the atom possesses an apparently infinite series of possible states of energies $\epsilon_1, \epsilon_2, \ldots, \epsilon_n, \ldots$, where

$$\epsilon_n = \frac{2\pi^2 e^4 m}{h^2}\left(1-\frac{1}{n^2}\right) = \chi\left(1-\frac{1}{n^2}\right), \qquad (7{\cdot}2)$$

and e is the electronic charge. The weights of these possible

* The assembly must contain molecules H_2 in various states as well as atoms, but they will be negligible in number at the temperatures which are of interest in applications, and we may confine ourselves to the simpler problem here.

† On the assumption that free ions or electrons behave as a perfect gas. This assumption is open to attack, but is presumably correct at sufficiently low pressures.

states must be determined on the principles laid down by Bohr; they are * given by the formula

$$p_n = n(n+1). \quad . \quad . \quad . \quad . \quad . \quad (7{\cdot}21)$$

Thus

$$b(\vartheta) = 2 + 6\vartheta^{\chi(1-2^{-2})} + 12\vartheta^{\chi(1-3^{-2})} + \ldots. \quad (7{\cdot}22)$$

This series, however, can never converge † for any value of ϑ, for $\epsilon_n \to \chi$ so that $\vartheta^{\epsilon_n} \to \vartheta^{\chi}$. The same difficulty would occur in discussing the reaction $X_2 \rightleftarrows 2X$ if we supposed that there might be an infinity of possible quantized states of axial vibration with energies less than χ.

This difficulty is not merely analytical but real and physical; for this whole statistical calculation (and all others ever made) are based on the hypothesis that the systems of the assembly, though they occasionally interact so that exchanges of energy are possible, yet for the greater part of time are pursuing their own undisturbed motions with a definite energy of their own. It is only so that there can be a definite energy distribution to study statistically. But the radius (mean distance) of the nth orbit is given by the formula

$$a_n = \frac{n^2h^2}{4\pi^2me^2}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (7{\cdot}3)$$

and tends to infinity with n. The full development, therefore, of a single H-atom may require all space; and it is not possible that for a given concentration the fundamental hypothesis can be obeyed in the case of these outer orbits. If owing to the physical conditions of the assembly these outer orbits cannot exist for any appreciable time or perhaps cannot exist at all, they may be ignored altogether in the discussion of the energy distribution.

Pure statistical theory knows no half-way house between an impossible state which can be ignored and a state which is to receive its full share of energy—calculated of course on quantum or classical principles as the case may be. It is unlikely that this represents physical facts. The distributions given by statistical theory can only be limiting cases, and it is to be presumed that in reality there can often be a continuous transition between them. This can be represented by "faking" the weights, but we are

* Bohr, 'The Quantum Theory of Line Spectra,' p. 76, footnote.

† Herzfeld (*loc. cit.*) has previously pointed out this difficulty. His discussion, however, hardly overlaps that of this section.

then introducing arguments foreign to the pure theory which we have tried to develop here. The difficulty is discussed more fully in the next section. It is sufficient to say here that some modification of $b(z)$ must be made, which on physical grounds must be equivalent to cutting off the series more or less sharply after a certain number of terms. Exactly where it should be cut off is difficult to decide.

If we insert known values in (7·1) we obtain, in terms of the absolute temperature T, the total pressure P, and the fraction of atoms ionized x,

$$\log \frac{x^2 P}{1-x^2} = -\frac{\chi}{kT} + 2{\cdot}5 \log T + \log \left\{ \frac{(2\pi m)^{3/2} k^{5/2}}{h^3} \right\} - \log B(T), \quad . \quad . \quad . \quad (7{\cdot}4)$$

which differs from Saha's formula in the term B(T) alone.

Saha's formula, which is obtained thermodynamically, may be regarded as derived on the assumption that the H-atom is either ionized or in its lowest quantum state. On his premises, therefore, our B(T) reduces to its first term p_1. If p_1 could be equal to 1, then (7·4) would be in agreement with Saha's formula, and the latter would be valid so long as the neglect of the other states of the H-atom is justifiable, which may sometimes or indeed generally be the case. But it does not appear possible on Bohr's principles to accept the value 1 for p_1. The true value is 2. The thermodynamic argument naturally, though of course not necessarily, forces in the value 1, as the idea of weight is somewhat foreign to it. We have here another example of the errors possible in a too uncritical use of the chemical constants.

In addition to a correction to Saha's formula for the error in p_1, the occurrence of the rest of B(T) (if any) must also be taken account of in applications of (7·4). The omission of this term by the thermodynamic argument is only apparent. It may be regarded thermodynamically as included in the specific heat term. But perhaps the present method has the advantage of putting it more clearly in evidence in proper relation to its origin ; in the thermodynamic argument it is natural to think of all the gases concerned as perfect, and therefore to assume $C_p = \frac{5}{2}R$, forgetting that the equation does not follow if the atoms can take up internal energy. The degree of ionization at given temperature and pressure may possibly be modified by the presence of B(T), which can be large at high temperatures if a large number of states are possible to

the H-atom. At the same time a large degree of ionization can already occur while the term $\chi/k\mathrm{T}$ is still fairly large; in such a case a large number of states must be admissible if this correction is to be a serious one. Further rediscussion of the ionization formula cannot be undertaken in detail here.

§ 8. *The temperature equilibrium of a gas of* H-*atoms.*— To reduce the difficulty due to the possible infinity of states of the H-atom to its simplest form, consider the temperature equilibrium of a gas of neutral H-atoms in which no ionization is supposed to occur. Let $h(z)$ and $b(z)$ be the partition functions for the translatory and internal energies. The state of zero energy may be taken to be that in which the atoms are all at rest, with their electrons in their lowest quantum orbits. If we suppose that the allowable states of the atom are cut off after N terms, all the ordinary arguments apply for any value of N to such an assembly. Then the total number of weighted complexions representing the assembly is

$$\mathrm{C} = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z^{\mathrm{E}+1}}[h(z)b(z)]^{\mathrm{X}}. \quad . \quad . \quad . \quad (8{\cdot}1)$$

If E_h and E_b are the mean values of the energy of translation and internal energy, then

$$\overline{\mathrm{E}_h} = \mathrm{X}\vartheta\frac{d}{d\vartheta}\log h(\vartheta) = \tfrac{3}{2}\mathrm{X}k\mathrm{T}, \quad . \quad . \quad . \quad . \quad (8{\cdot}2)$$

$$\overline{\mathrm{E}_b} = \mathrm{X}\vartheta\frac{d}{d\vartheta}\log b(\vartheta), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8{\cdot}21)$$

and $\mathrm{E} = \overline{\mathrm{E}_h} + \overline{\mathrm{E}_b}$.

Moreover,

$$b(\vartheta) = 2 + 6\vartheta^{\chi(1-2^{-2})} + \ldots + \mathrm{N}(\mathrm{N}+1)\vartheta^{\chi(1-\mathrm{N}^{-2})}, \quad . \quad (8{\cdot}3)$$

$$\vartheta b'(\vartheta) = \chi\{6(1-2^{-2})\vartheta^{\chi(1-2^{-2})} + \ldots + \mathrm{N}(\mathrm{N}+1)(1-\mathrm{N}^{-2})\vartheta^{\chi(1-\mathrm{N}^{-2})}\}. \quad (8{\cdot}31)$$

Now if ϑ has any fixed value whatever, and we let $\mathrm{N}\to\infty$, it is easy to see that

$$\frac{\vartheta b'(\vartheta)}{b(\vartheta)} \to \chi. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8{\cdot}4)$$

In spite, therefore, of the non-convergence of $b(\vartheta)$, we can

arrive at a definite value for $\overline{E_b}$ without restriction to the number of terms; we have, in fact,

$$\overline{E_b} = X\chi. \quad . \quad . \quad . \quad . \quad . \quad . \quad (8{\cdot}41)$$

This equality must hold in the limit for all values of ϑ (or T) except $\vartheta=0$, when we have obviously $\overline{E_b}=0$.

Consider more closely the implications of such a result. Given E, the temperature is to be determined by the equation

$$E = X\chi + \tfrac{3}{2}XkT.$$

This equation is impossible of fulfilment unless $E > X\chi$. The impossibility has of course entered in the limiting process $N\to\infty$. The equation must really mean that all the energy goes into lifting the electrons into orbits of high quantum numbers, the temperature remaining practically zero until enough energy ($X\chi$) has been supplied practically to ionize every atom. After that, on a further supply of energy, the usual kinetic energy of translation is built up with these already satiated atoms, no further internal energy being required. There does not seem to be any escape from these conclusions if the preceding analysis could really be pushed home without effective limitation on the possible states. In fact (granted the hypotheses) these conclusions seem to be in agreement with the dictates of common sense.

These conclusions of course cannot stand, and therefore restrictions on the possible states of the H-atoms are essential. We have already pointed out that the mean distance of the atoms provides such a limitation; for atoms of the size implied by the extended orbits must be able to exist sufficiently long for the equilibrium state to be set up. The tendency of the atom to radiate spontaneously, or under the stimulus of radiation, may also impose an additional control, which effectively cuts out high quantum number orbits from the statistical equilibrium.

§ 9. *General dissociation equilibria.*—We now return to the general dissociation problem, conforming as closely as possible to the treatment of Ehrenfest and Trkal. Consider an assembly of s different kinds of atoms, in number $X_1, \ldots, X_s$, capable of forming j different kinds of molecules. In any example of the assembly let us suppose that there are $M_1, \ldots, M_s$ free atoms of each sort and $N_1, \ldots N_j$ of each of the different molecules. Suppose in addition that the rth type of molecule, for example,

is composed of $q_{1r}, q_{2r}, \ldots, q_{sr}$ atoms of the various types $1, \ldots, s$, the q's being all positive integers. Then

$$M_t + \Sigma_r q_{tr} N_r = X_t, \quad (t = 1, \ldots, s). \quad . \quad . \quad (9{\cdot}1)$$

Let the state of zero energy of the assembly be defined as that state in which all the molecules are dissociated and all the resulting atoms at rest. For a single molecule the state of zero energy is taken to be that of no motion of translation, no rotation, or the rotational quantum numbers zero, and no internal vibrations of any type, or all the vibrational quantum numbers zero *. Let χ_r be the heat of dissociation for the molecule of type r—that is, the constant work required to dissolve such a molecule at rest in its lowest quantum state into its separate component atoms also at rest. Let the partition functions be $g_1(z), \ldots, g_s(z)$, $h_1(z), \ldots, h_j(z)$ for the translational motion of the atoms and molecules, and $b_1(z), \ldots, b_j(z)$ for the rotations and internal vibrations of the molecules: it is assumed that the series for the latter either converge or terminate. The symmetry numbers of the molecules are $\sigma_1, \ldots, \sigma$. If the total energy of the assembly is E, and F is the energy for distribution by the partition functions when there exist $N_1, \ldots, N_j$ molecules of the various types, then

$$F = E + \Sigma_r N_r \chi_r. \quad . \quad . \quad . \quad . \quad . \quad . \quad (9{\cdot}2)$$

For any example of the assembly with $M_1, \ldots, M_s$ atoms and $N_1, \ldots, N_j$ molecules the number of weighted complexions c is given by

$$c = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z^{F+1}} [g_1(z)]^{M_1} \ldots [g_s(z)]^{M_s} [h_1(z)\, b_1(z)]^{N_1} \ldots \times [h_j(z)\, b_j(z)]^{N_j}. \quad . \quad (9{\cdot}3)$$

The number of different examples of such an assembly is †

$$\frac{X_1! \ldots X_s!}{M_1! \ldots M_s!\ N_1! \ldots N_j!\ \sigma_1^{N_1} \ldots \sigma_j^{N_j}}, \quad . \quad . \quad (9{\cdot}31)$$

and therefore the total number of weighted complexions representing the assembly is

$$C = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z^{E+1}} \Sigma_{(N)} \frac{X_1! \ldots X_s!}{M_1! \ldots M_s!\ N_1! \ldots N_j!} [g_1(z)]^{M_1} \ldots [g_s(z)]^{M_s} \times \left[\frac{h_1(z) b_1(z)}{\sigma_1 z^{\chi_1}}\right]^{N_1} \ldots \left[\frac{h_j(z) b_j(z)}{\sigma_j z^{\chi_j}}\right]^{N_j}. \quad (9{\cdot}32)$$

* As in § 4, it is easy to make the necessary modifications for the case in which zero values for any or all of these quantum numbers are to be excluded.

† Ehrenfest and Trkal, *loc. cit.* p. 169.

Using the equations (9·1), we can reduce this to the form

$$C = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z^{E+1}}[g_1(z)]^{X_1}\dots[g_s(z)]^{X_s} \times \Sigma_{(N)}\frac{X_1!\dots X_s!\,\beta_1^{N_1}\dots\beta_j^{N_j}}{M_1!\dots M_s!\,N_1\dots N_j!}, \quad (9\cdot33)$$

where

$$\beta_r = \frac{h_r(z)b_r(z)}{\sigma_r z^{X_r}\{g_1(z)\}^{q_{1r}}\dots\{g_s(z)\}^{q_{sr}}}, \quad (r=1,\dots,j). \quad (9\cdot34)$$

The summation $\Sigma_{(N)}$ is to be taken over all possible values of $N_1,\dots,N_j$—that is, over all zero and positive integral values which are small enough to leave the M's positive or zero. In the same way,

$$\overline{CN_r} = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z^{E+1}}[g_1(z)]^{X_1}\dots[g_s(z)]^{X_s} \times \Sigma_{(N)}\frac{N_r X_1!\dots X_s!\,\beta_1^{N_1}\dots\beta_j^{N_j}}{M_1!\dots M_s!\,N_1!\dots N_j!}, \quad (r=1,\dots,j). \quad (9\cdot35)$$

The series in (9·35) is obviously equal to

$$\beta_r\frac{\partial}{\partial\beta_r}\Sigma_{(N)},$$

where $\Sigma_{(N)}$ is the series in (9·33), the equation for C.

To evaluate the integrals for C and CN_r we proceed as in § 5. By the arguments there used it follows from (9·32) that the integrand on any circle γ has its maximum on the real axis, and on the real axis has a unique minimum between 0 and 1 at $z=\vartheta$, where ϑ is the sole root of

$$E = \Sigma_t X_t\vartheta\frac{d}{d\vartheta}\log g_t(\vartheta) + \vartheta\Sigma_r\left\{\frac{d\beta_r}{d\vartheta}\frac{\partial}{\partial\beta_r}\log\Sigma_{(N)}\right\}. \quad (9\cdot4)$$

The discussion reduces as before to a study of the function $\log\Sigma_{(N)}$, which can be shown by an obvious extension of the former argument to possess a unique maximum term for values of the N's determined by the equations

$$N_r/(M_1^{q_{1r}}M_2^{q_{2r}}\dots M_s^{q_{sr}}) = \beta_r, \quad (r=1,\dots,j). \quad (9\cdot5)$$

These equations, j in number, with the s equations (9·1), are sufficient to determine the $(s+j)$ unknowns (M_t, N_r). In the neighbourhood of the maximum term (M_t^0, N_r^0) the

terms of the series can be expressed to a sufficient approximation in the form

$$\frac{1}{(2\pi)^{\frac{1}{2}j}}\left(\frac{X_1 \ldots X_s}{M_1^0 \ldots M_s^0\, N_1^0 \ldots N_j^0}\right)^{\frac{1}{2}} \exp\left\{-\tfrac{1}{2}\Sigma_r \frac{\alpha_r^2}{N_r^0}\right.$$
$$-\tfrac{1}{2}\Sigma_t \frac{(\Sigma_r q_{tr}\alpha_r)^2}{M_t^0} + \Sigma_r N_r^0 \log \beta_r + \Sigma_t (X_t \log X - X_t)$$
$$\left.-\Sigma_r(N_r^0 \log N_r^0 - N_r^0) - \Sigma_t(M_t^0 \log M_t^0 - M_t^0)\right\}. \quad . \quad . \quad . \quad (9{\cdot}51)$$

We can then, instead of summing the series, integrate (9·51) j times with respect to the α's, and obtain with sufficient accuracy

$$\log \Sigma_{(N)} = \Sigma_r N_r^0 \log \beta_r - \Sigma_t(X_t \log X_t - X_t)$$
$$-\Sigma_r(N_r^0 \log N_r^0 - N_r^0) - \Sigma_t(M_t^0 \log M_t^0 - M_t^0). \quad . \quad . \quad . \quad (9{\cdot}52)$$

For the series in $\overline{N}_r$ we obtain the same value as for $\Sigma_{(N)}$ multiplied by N_r^0. If, further, we differentiate (9·52) partially with respect to β_1 say, we find

$$\beta_1 \frac{\partial}{\partial \beta_1} \log \Sigma_{(N)} = N_1^0 + \Sigma_r \beta_1 \frac{\partial N_r^0}{\partial \beta_1}(\log \beta_r - \log N_r^0)$$
$$+\Sigma_t \beta_1 \frac{\partial M_t^0}{\partial \beta_1} \log M_t^0.$$

If we express the $\partial M_t^0/\partial \beta_1$ in terms of the $\partial N_r^0/\partial \beta_1$ by using (9·1) we obtain coefficients of $\partial N_r^0/\partial \beta_1$ which all vanish in virtue of (9·5). Thus

$$\beta_1 \frac{\partial}{\partial \beta_1} \Sigma_{(N)} = N_1^0 \Sigma_{(N)}, \quad . \quad . \quad . \quad . \quad (9{\cdot}6)$$

and in general

$$\beta_r \frac{\partial}{\partial \beta_r} \Sigma_{(N)} = N_r^0 \Sigma_{(N)}, \quad (r = 1, \ldots, j). \quad . \quad (9{\cdot}61)$$

This is in agreement with the result above, so that (9·52) can be differentiated with respect to any of the β's and therefore with respect to z.

We have now all the necessary machinery to evaluate the partition, and find at once that $\overline{N}_r = N_r^0$, $(r = 1, \ldots, j)$, provided that the β's are evaluated for $z = \vartheta$. Thus the dissociation-equilibria must be determined by the j "laws

of mass-action" in the form

$$\overline{N}_r/\{(\overline{M}_1)^{q_{1r}}\dots(\overline{M}_s)^{q_{sr}}\} = \frac{h_r(\vartheta)\,b_r(\vartheta)}{\sigma_r\vartheta^{\chi_r}\{g_1(\vartheta)\}^{q_{1r}}\dots\{g_s(\vartheta)\}^{q_{sr}}},$$
$$(r=1,\dots,j). \quad (9{\cdot}7)$$

Equation (9·4) relating the temperature and energy reduces to

$$E = \Sigma_t\overline{M}_t\vartheta\frac{d}{d\vartheta}\log g_t(\vartheta) + \Sigma_r\overline{N}_r\vartheta\frac{d}{d\vartheta}\{\log h_r(\vartheta)+\log b_r(\vartheta)\} - \Sigma_r\overline{N}_r\chi_r, \quad (9{\cdot}71)$$

and the further details of the partition can be worked out exactly as in § 6. The result, that the details are the same as if the M's and N's were fixed at their equilibrium values, still holds.

It is convenient to express (9·7) in a form more in accord with the equivalent thermodynamical formulæ. We have for the whole assembly

$$PV = kT(\Sigma_t\overline{M}_t+\Sigma_r\overline{N}_r), \quad (9{\cdot}8)$$

where P is the total pressure. If x_t, x_t^* are the "fractional concentrations" † of the various atoms and molecules respectively, then

$$x_t = \frac{\overline{M}_t}{\Sigma\overline{M}_t+\Sigma_r\overline{N}_r}, \quad x_r^* = \frac{\overline{N}_r}{\Sigma_t\overline{M}_t+\Sigma_r\overline{N}_r}. \quad (9{\cdot}81)$$

We then express (9·7) in logarithmic form, eliminate $\overline{M}_t$, $\overline{N}_r$, and V by these expressions, and insert known values for $g_t(\vartheta)$, $h_r(\vartheta)$, and ϑ. We obtain after reducing the equations

$$\Sigma_t q_{tr}\log x_t - \log x_r^* = (\Sigma_t q_{tr}-1)(-\log P+\tfrac{5}{2}\log T) - \frac{\chi_r}{kT} - \log B_r(T) + \Sigma_t q_{tr}\log\left\{\frac{(2\pi m_t)^{3/2}k^{5/2}}{h^3}\right\} - \log\left\{\frac{(2\pi m_r^*)^{3/2}k^{5/2}}{\sigma_r h^3}\right\}, \quad (r=1,\dots,j). \quad (9{\cdot}82)$$

In this equation $B_r(T)$ is the rotational and vibrational partition function expressed in terms of T, and m_t, m_r^* are the

† "Concentrations" is the term used by Ehrenfest and Trkal in this connexion. It is, however, not here used in its natural meaning of mass or number of molecules per c.c. Equation (9·82) is the reaction-isobar expressed effectively in terms of *partial pressures*.

masses of the atoms and molecules respectively. There may be, of course, a very large number of other mass-action equations among the constituents of the assembly, but they must all be expressible in terms of the j fundamental mass-action equations (9·7) or (9·82), in each of which only one complex constituent occurs. It is therefore unnecessary to consider such other equations in detail here.

In the cases considered by Ehrenfest and Trkal, $B_r(T)$ has its limiting form to which it approximates when the rotations are fully excited and the internal vibrations completely unexcited. In such cases (9·82) agrees with their results *.

Finally it is interesting to verify that the form of the reaction-isobar thus derived agrees with that derived by thermodynamic arguments from van't Hoff's equation,

$$\frac{d\log K}{dT} = -\frac{Q_p}{RT^2}, \quad . \quad . \quad . \quad . \quad . \quad (9·9)$$

where Q_p is the heat of reaction per gram-molecule at constant pressure and K the equilibrium constant; in particular, we shall see how the term B(T) enters the reaction-isobar. By Kirchhoff's equation,

$$\frac{dQ_p}{dT} = -\Sigma q C_p, \quad . \quad . \quad . \quad . \quad . \quad (9·91)$$

where C_p is the specific heat at constant pressure, the summation is taken over all the reacting species, and q is the number of gram-molecules of each species with a negative sign for those that disappear. Therefore

$$\begin{aligned} Q_p &= Q_0 - \int_0^T \Sigma q C_p dT = Q_0 - \Sigma q RT - \int_0^T \Sigma q C_v dT, \\ &= Q_0 - \Sigma q RT - \Sigma q (U_v - (U_v)_0), \quad . \quad . \quad . \quad . \quad (9·92) \end{aligned}$$

where Q_0 is the limit of the heat of reaction as $T \to 0$ and U_v is the internal energy per gram-molecule at volume v. Therefore, on substituting in (9·9) and integrating, we have

$$\log K = \frac{Q_0}{RT} + \Sigma q \log T + \int_0^T \frac{dT}{RT^2} \Sigma q (U_v - (U_v)_0) + A, \quad (9·93)$$

where A is a constant. Equation (9·93) is the general thermodynamic formula. We can, however, translate it at once into our notation, and it then integrates in finite

* *Loc. cit.*, equation (41).

terms. For

$$Q_0/R = -\chi/k,$$

$$U_v - (U_v)_0 = N\vartheta \frac{d}{d\vartheta}\{\log h(\vartheta) + \log b(\vartheta)\},$$

$$\frac{dT}{RT^2} = \frac{1}{N}\frac{d\vartheta}{\vartheta},$$

where N is Loschmidt's number. Thus

$$\log K = -\frac{\chi}{kT} + \Sigma q \log T + \Sigma q \{\log h(\vartheta) + \log b(\vartheta)\} + A. \quad . \quad . \quad . \quad (9{\cdot}94)$$

Equations (9·94) and (9·82) are equivalent, except that in (9·82) the constant A is determined.

§ 10. *The vapour-pressure equation.*—We can derive in an exactly similar way the equation for the vapour-pressure of a solid—in particular, of a crystal. The contributions to the equation of the internal energy of the vapour molecule and the internal energy of the crystal can be formally included in the statistical discussion. It is not possible to include here a proper discussion of the partition function for a crystal without unduly lengthening this paper. This must be discussed elsewhere *. It is sufficient to point out that for a crystal of N molecules a partition function $[K(z)]^N$ can be constructed, based on Born and Karman's † results.

Following Ehrenfest and Trkal, we suppose that there are in the assembly N free molecules and a single crystal of M molecules. The molecules are composed out of s types of atom, $X_1, \ldots, X_s$ in number, and there is a total of X molecules in all. The problem is actually one of dissociation-equilibrium in which there are only two kinds of molecule present—the actual free molecules N in number, and the whole crystal of which there is 1 ‡.

Let us suppose for definiteness that the assembly has zero energy when all the molecules are condensed on the crystal, and all the internal motions of the crystal vanish. We take our usual definition of zero energy for a free molecule, and suppose that χ is the work necessary to remove one molecule from the crystal when both crystal and molecule have (by convention) zero energy. It is easy to make the necessary modifications

* Darwin and Fowler, Proc. Camb. Phil. Soc. (1922).

† Born, *Der Dynamik der Krystallgitter*, p. 72.

‡ Or at least a number negligible compared with X.

if any of the corresponding quantum numbers cannot have the value zero. Then for any example of the assembly with N free molecules we have in the usual notation

$$\mathrm{F} = \mathrm{E} - \mathrm{N}\chi, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10{\cdot}1)$$

$$c = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z^{\mathrm{F}+1}}[\mathrm{K}(z)]^{\mathrm{M}}[b(z)h(z)]^{\mathrm{N}}, \quad (10{\cdot}11)$$

$$\mathrm{N} + \mathrm{M} = \mathrm{X}.$$

The number of such examples of the assembly is

$$\frac{\mathrm{X}_1 ! \ldots \mathrm{X}_s !}{\mathrm{N} !\, \sigma^{\mathrm{N}}\, 1 !\, \sigma'}, \quad . \quad . \quad . \quad . \quad . \quad (10{\cdot}12)$$

where σ is the symmetry number of the free molecule and σ' the symmetry number of the crystal. Strictly speaking, there should also be included in (10·11) a partition function $\mathrm{H}(z)$ for the mass motion of the whole crystal or crystals (Brownian movement), but this can obviously be ignored. The constant factor $(\mathrm{X}_1 ! \ldots \mathrm{X}_s !)/\sigma'$ in (10·12) can also be dropped out. With these abbreviations, we have

$$\mathrm{C} = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z^{\mathrm{E}+1}}[\mathrm{K}(z)]^{\mathrm{X}}\Sigma_{\mathrm{N}}\frac{1}{\mathrm{N}!}\left[\frac{b(z)h(z)z^{\chi}}{\sigma \mathrm{K}(z)}\right]^{\mathrm{N}}, \quad (10{\cdot}2)$$

in which the summation goes from 0 to X. If X is taken sufficiently large, it is easy to show as usual that Σ_{N} can be replaced for all purposes by its limit, which here has the simple form

$$\exp\left\{\frac{b(z)h(z)z^{\chi}}{\sigma \mathrm{K}(z)}\right\}. \quad . \quad . \quad . \quad . \quad (10{\cdot}21)$$

In the same way

$$\mathrm{C}\overline{\mathrm{N}} = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z^{\mathrm{E}+1}}[\mathrm{K}(z)]^{\mathrm{X}}\left\{\frac{b(z)h(z)z^{\chi}}{\sigma \mathrm{K}(z)}\right\}\exp\left\{\frac{b(z)h(z)z^{\chi}}{\sigma \mathrm{K}(z)}\right\},$$

and it follows at once that . . . (10·22)

$$\overline{\mathrm{N}} = \frac{b(\vartheta)h(\vartheta)\vartheta^{\chi}}{\sigma \mathrm{K}(\vartheta)}, \quad . \quad . \quad . \quad . \quad (10{\cdot}3)$$

which is the vapour-pressure equation. In its more usual form it reads

$$\log p = -\frac{\chi}{k\mathrm{T}} + \frac{5}{2}\log \mathrm{T} + \log\left\{\frac{(2\pi m)^{3/2}k^{5/2}}{\sigma h^3}\right\} + \log \mathrm{B}(\mathrm{T}) - \log \mathrm{K}(\mathrm{T}). \quad (10{\cdot}31)$$

Ehrenfest and Trkal * give (10·31) in the form to which it reduces when $K(T)=1$, so that the energy of the crystal is ignored and $B(T)$ is replaced by its limit when all three rotations are excited, while all the internal vibrations are at rest.

§ 11. *The chemical constant.*—Equation (10·31) is the most satisfactory starting-point for a discussion of the chemical constant. The chemical constant may be defined † as the constant i of the vapour-pressure equation in the form

$$\log p = -\frac{\chi}{kT}+\frac{5}{2}\log T+i, \quad . \quad . \quad (11\cdot11)$$

for a monatomic gas or one which so behaves at the temperatures considered (H_2 at very low temperatures). For a diatomic gas with rotations fully excited, the equation becomes

$$\log p = -\frac{\chi}{kT}+\frac{7}{2}\log T+i', \quad . \quad . \quad (11\cdot12)$$

and for a triatomic gas with all three rotations excited

$$\log p = -\frac{\chi}{kT}+4\log T+i''. \quad . \quad . \quad (11\cdot13)$$

The heat content of the solid is neglected throughout for simplicity of statement ; it can be included if desired.

If we apply the exact equation (10·31), putting $K(T)=1$, we can at once assign the theoretical value of i. The interesting case is (11·11). If the gas is genuinely monatomic, then $B(T)=1$, $\sigma=1$, and

$$i = \log \frac{(2\pi m)^{3/2}k^{5/2}}{h^3}, \quad . \quad . \quad . \quad . \quad (11\cdot2)$$

which is the usual well-known formula ‡. Suppose next that the gas is diatomic, and in the region in which the rotations are practically unexcited. Then for such a gas as H_2, $\sigma=2$. If the state of no rotation is possible, then $B(T)\rightarrow 1$ and

$$i = \log \frac{(2\pi m)^{3/2}k^{5/2}}{2h^3}; \quad . \quad . \quad . \quad (11\cdot21)$$

* *Loc. cit.*, equations (85) and (46).

† Nernst, *Die theoretischen und experimentellen Grundlagen des neuen Wärmesatzes*, p. 136 (1918). Allowances, of course, can and should be made for the other terms varying with T which tend to zero as $T\rightarrow 0$.

‡ Nernst, *loc. cit.* p. 136.

if the state of no rotation is impossible, then

$$\mathrm{B(T)} \backsim p_1 e^{-\epsilon/k\mathrm{T}}.$$

The term $\epsilon/k\mathrm{T}$ must, as we saw in § 4, be absorbed in χ, and we are left with

$$i = \log \frac{p_1(2\pi m)^{3/2}k^{5/2}}{2h^3}. \quad . \quad . \quad . \quad (11{\cdot}22)$$

In such a case, therefore, the chemical constant may depend not only on the mass but also on the symmetry number * *and the weight of the lowest quantum state of the molecule.*

But we can draw certain interesting conclusions from the experimental data for H_2 in comparison with genuinely monatomic vapours. It is found † that the chemical constants for mercury, argon, zinc, cadmium, and hydrogen all agree within experimental error with the theoretical value calculated by formula (11·2). The fact that this formula holds for H_2 can only mean that formula (11·22) must be used with $p_1=2$. The inclusion of a factor $\frac{1}{2}$ or $\frac{3}{2}$ in the logarithm brings the value of i well outside the permissible range of error. We have, therefore, strong additional evidence (1) that the state of no rotation is impossible, and (2) that the weight of the one-quantum state of rotation is 2 and not 1 or 3. It is interesting to remark that it is only by a fortuitous cancelling of two errors that the formula for a monatomic gas is applicable to H_2 at all. It would be exceedingly interesting if the chemical constant of some non-symmetrical diatomic molecule could be determined in a region in which the rotations are practically unexcited. For such a molecule $\sigma=1$, and the chemical constant should be greater by log 2 than the value derived from (11·2).

The theoretical values of i' and i'' are also given by Ehrenfest and Trkal. The value of i' can be obtained from the limiting form of B(T) when the rotations are fully excited and the axial vibrations unexcited (see § 4). Thus

$$\mathrm{B(T)} \backsim \frac{8\pi^2 \mathrm{J}k\mathrm{T}}{h^2},$$

$$i' = \log \frac{4\pi(2\pi m)^{3/2}(2\pi \mathrm{J})k^{7/2}}{\sigma h^5}, \quad . \quad . \quad (11{\cdot}3)$$

* The dependence on the symmetry number is pointed out by Ehrenfest and Trkal, *loc. cit.* p. 164.

† Nernst, *loc. cit.* p. 151; Egerton, Phil. Mag. vol. xxxix. p. 15 (1920).

which agrees with their value *. Agreement, however, is impossible if the weights n or $n+1$ are taken for the nth quantum state of the rotational motion. Such weights are therefore inadmissible, and there seems little doubt, in view of the additional evidence presented here, that $2n$ is the correct theoretical value of the weight of the nth quantum state for the rotations of a diatomic molecule. This molecule also, on the whole, fits best the demands of the curve of specific heats of H_2 †; but in this connexion one cannot discriminate between the weights $2n$ and n. The corresponding general value of i'' cannot at present be obtained from the limiting form of B(T), for the quantized rotations of a rigid body with three unequal moments of inertia have not yet been completely worked out. It can, however, be obtained by treating the rotations throughout as obeying classical laws, as is done by Ehrenfest and Trkal. Its value is

$$i'' = \log \frac{8\pi^2(2\pi m)^{3/2}(8\pi^3 J_1 J_2 J_3)^{1/2} k^4}{\sigma h^6}, \quad . \quad . \quad (11{\cdot}4)$$

where J_1, J_2, J_3 are the moments of inertia.

Returning to the dissociation equation (9·82), we see that the constant term in this equation is composed of the same chemical constants that occur in the vapour-pressure equations, and that no new constant enters. Equation (9·82) can be written, when the internal vibrations of the molecule are ignored and its rotations are either completely excited or unexcited,

$$\Sigma_t q_{tr} \log x_t - \log x_r^* = -(\Sigma_t q_{tr} - 1)\log P + \alpha \log T + \Sigma_t q_{tr} i_t - i_r^*, \quad (r=1, \ldots, j). \quad (11{\cdot}5)$$

In this equation α depends on how many rotations of the molecule are fully excited, and i_t, i_r^* are the chemical constants, according to the equations of this section, of the atoms and molecules respectively.

This brings to a conclusion the statistical theory of ordinary dissociation-equilibria. The theorems proved, which were summarized in § 1, have obviously a bearing on general thermodynamics, but this development cannot be undertaken here.

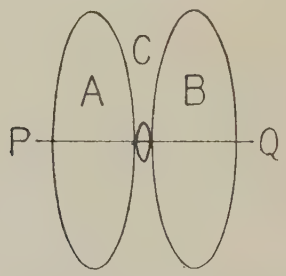
Cambridge, Aug. 1922.

* *Loc. cit.*, equation (46). † Reiche, *loc. cit.*

II. *The Process of Light Radiation from an Atom of a Gas leading to a Possible Interpretation of Planck's Quantum by means of Electrodynamics.* IV. *By* Albert C. Crehore *.

SUPPOSE that the atom of fig. 1, which represents a section of a normal hydrogen atom as described in the first paper †, is one of the atoms in a mass of hydrogen gas in a vacuum tube in its steady state not emitting any characteristic light radiation, and that this gas is suddenly bombarded by a stream of negative electrons projected through it. Let us endeavour to trace the effect that one of these negative electrons passing near to a hydrogen atom of this kind, but not colliding with it, will have upon the atom merely because of the mutual dynamical action between the two bodies. The total effect may be studied by considering

Fig. 1.



Meridian section of the hydrogen atom described in the first paper.

independently the effects that the passing electron exerts upon the three individual charges making up the atom by means of the general expression for the ponderomotive force due to Saha that a second point charge in motion (the passing electron) exerts upon the first point charge (the separate individuals in the hydrogen atom). This force has been given in the first paper ‡ as

$$\mathbf{F} = \frac{E_1 E_2 (1-\beta_2^2)}{k(RA_2)^3 (1-\beta_1^2)^{\frac{1}{2}}} \left\{ \left(1 - \frac{1}{c^2}\mathbf{q}_1 \cdot \mathbf{q}_2\right)\mathbf{R} - \frac{RA_1}{c}\mathbf{q}_2 \right\}, \quad . \quad (1)$$

* Communicated by the Author.

† Phil. Mag. Oct. 1921, p. 588.

‡ *Loc. cit.* Eq. (3), p. 570. This should read as in (1) above. An obvious misprint is the omission of the exponent 2 from the c in the term $\mathbf{q}_1 \cdot \mathbf{q}_2/c^2$.

where the subscripts 1 and 2 refer to the first and second point charges respectively. The letters A_1 and A_2 are abbreviations for the expression $1-\frac{\mathbf{q}\cdot\mathbf{R}}{c\mathrm{R}}$ for the first and second points respectively.

For clearness of presentation let us at first neglect the fact that each individual body of charge is supposed to be a rotating body, and treat each simply as a static charge. Considering the positive charge of the atom ($+2e=\mathrm{E}_1$) as a body at rest upon which we are to express the force due to the passing negative electron $-e$, we have

$$\mathfrak{q}_1=0,\ \beta_1=0,\ \mathrm{A}_1=1,\ \text{and}\ \mathrm{E}_1\mathrm{E}_2=-2e^2.$$

Hence (1) reduces to

$$\mathbf{F}=-\frac{2e^2(1-\beta_2^2)}{\mathrm{k}(\mathrm{RA}_2)^3}\left\{\mathbf{R}-\frac{\mathrm{R}}{c}\mathfrak{q}_2\right\}, \quad . \quad . \quad . \quad (2)$$

An estimate of the probable value of the order of magnitude of $\mathfrak{q}_2$, the velocity of the passing electron, may be obtained from the experimental observation that light radiation does not begin in hydrogen gas until the velocity of the passing electron corresponds to about 11 volts, that is to say a velocity of about 2×10^8 cm./sec., or $\beta_2=\frac{2}{3}\times10^{-2}$, less than the one hundredth part of the velocity of light. At this small velocity the part of the force expressed by the second term of (2) has less than one hundredth of the value of the part expressed by the first term, and may for a first approximation be neglected. If, therefore, we put also

$$\mathfrak{q}_2=0,\ \beta_2=0,\ \text{and}\ \mathrm{A}_2=1,\ (2)\ \text{reduces to}$$

$$\mathbf{F}=-\frac{2e^2}{\mathrm{kR}^3}\mathbf{R}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

which is merely the electrostatic force, this being an attraction, since by definition the vector $\mathbf{R}$ has its positive direction from the second to the first point charge.

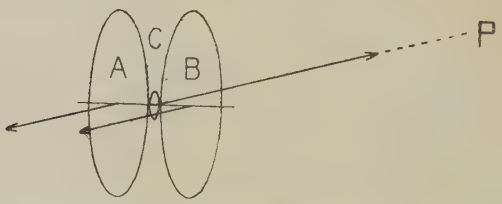
In a similar manner the force upon each individual negative electron of the atom becomes

$$\mathrm{F}=\frac{e^2}{\mathrm{kR}^3}\mathbf{R}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

a repulsion of half the value of (3), being in the opposite

direction. These forces are represented in fig. 2 by vectors drawn from the centre of each charge of the atom parallel to **R**, that is to the line from the distant passing electron P to the atom. The sum of these forces is zero as far as any translational effect upon the atom as a whole is concerned, considering the three R's to be approximately the same, but there is produced an internal effect within the atom itself. In general, unless P happens to be exactly in the equatorial plane of the atom, there is a component of each force which acts in the direction of its axis. The one pair of charges B and C, are held by this component more firmly together, while the other pair, A and C, are less firmly held and tend to separate. The chief effect of this component is, therefore, to increase the momentary pressure between B and C

Fig. 2.



Illustrating the forces upon the three parts of the atom due to the passing electron, P, electrostatic terms only.

and to decrease that between A and C, thus rendering A more readily detachable from the atom than B when other effects arise from any cause.

The other component of the forces resolved in the equatorial plane produces a couple tending to displace the two electrons A and B in the same direction off from the axis of C, but these electrostatic forces are not great enough to cause much displacement, and cannot be until the passing electron is near enough to be in contact with the atom. This near approach is not necessary to account for radiation from the atom, as will be shown.

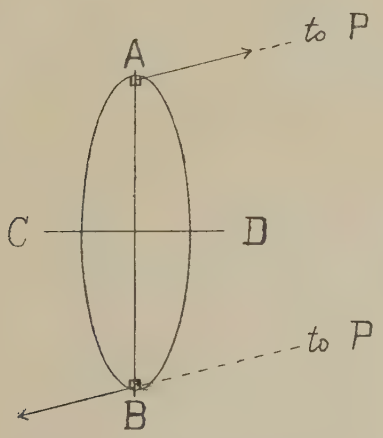
The rotation upon their axes of the charges A, B, and C having a non-spherical shape is the key to the behaviour of this atom. Let us next consider the effect that the passing electron P must have upon the positive charge C as a rotating charge. In fig. 3 the ellipse represents a meridian section of the positive charge, and P the distant passing

electron. Apply the formula (1) to find the force upon an element ($d\mathrm{E}_1$) of the positive charge in its upper side at A due to the passing electron of charge $-e_2$ as point charges. For small velocities of P of the order of, say, the 1/100 the velocity of light, the quantities A_1 and A_2 in the general formula are approximately unity, and (1) becomes

$$\mathbf{F} = -\frac{e_2 d\mathrm{E}_1}{\mathrm{k}\mathrm{R}^3}\left\{\mathbf{R} - \frac{\mathrm{R}}{c}\mathbf{q}_2 - \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{c^2}\mathbf{R}\right\}. \quad . \quad . \quad (5)$$

The first term within the brace, $\mathbf{R}$, is the electrostatic force discussed above acting along the line $\mathbf{R}$. The term $-\frac{\mathrm{R}}{c}\mathbf{q}_2$ is

Fig. 3.



Illustrating the couple acting upon a pair of elements A and B of the positive charge due to the passing electron P originating in the term $(1/c^2)$ $(\mathbf{q}_1 \cdot \mathbf{q}_2)\mathbf{R}$, which tends to rotate the axis CD.

small by comparison when β_2 is less than 1/100, and has a direction the same as $\mathbf{q}_2$, the velocity of P. Since the velocity, $\mathbf{q}_1$, of the first point A does not enter this expression, the force due to this term is substantially the same upon every element of the positive charge. This term, therefore, has the same kind of effect upon the positive charge as has the electrostatic force, only much smaller and in a different direction, so that the first two terms of (5) may be classed together and dismissed from further consideration, as merely producing a slight change in pressure between the electrons A and B upon C of the model (fig. 1).

The remaining term of (5) becomes the important one, and the effective force due to this term upon the element at A (fig. 3) due to P to produce a rotation of the axis CD, becomes

$$\mathbf{F} = -\frac{e_2 d\mathrm{E}_1}{\mathrm{kR}^3}\left\{-\frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{c^2}\mathbf{R}\right\}, \quad . \quad . \quad . \quad . \quad (6)$$

which is represented in the figure by the vector from A towards P, taking the direction of $-\mathbf{R}$. When the element A is moving either from or towards the observer perpendicular to the paper, then the motion of the element B diametrically opposite is in exactly the reverse direction. For this element the cosine of the angle between $\mathbf{q}_1$ and $\mathbf{q}_2$ has the opposite sign, and the force upon the element B has the opposite direction along $\mathbf{R}$, represented by the equal vector from B away from P. This pair of forces upon the two elements forms a couple having a tendency to turn the plane of rotation of the positive charge unless P happens to be exactly in the plane of the equator as a special case. The total effect of this term in the force equation (6) upon the positive charge at any moment is, of course, the integral of these elementary forces throughout its volume. Without undertaking to integrate it, however, the chief conclusion is that there is no doubt that the total effect of the motion of P is to produce in the positive charge a tendency to turn its axis of rotation CD.

There exists at the same time a similar effect of the passing electron upon each negative electron A and B in the atom; but it may be shown that the angular acceleration produced in these electrons is small in comparison with that produced in the positive charge, and therefore for the present purpose it is negligible.

To compare their orders of magnitude, consider a pair of elements of charge as before, and take the position of P on the axis so that the vector $\mathbf{R}$ is perpendicular to the plane of the orbit, and the effect is a maximum. The moment of the couple for two opposite elements, A and B of fig. 3, of the positive charge is then $2a\mathrm{F}$, where a is the radius or distance from the element to the axis CD, and F the force upon it as expressed by (6). Setting this equal to the product of its moment of inertia I by the angular acceleration of the axis α, we find the angular acceleration in terms of the element of mass dm to be

$$\alpha = \frac{2a\mathrm{F}}{\mathrm{I}} = \frac{2\mathrm{F}}{a\,dm}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

Next suppose the negative electron to be divided up into the same number of elements as the positive charge, and select a pair of elements similarly situated in the oblate spheroid to those chosen in the positive charge. The elementary positive charge will then be twice as great as the elementary negative charge, because the whole positive charge is twice the negative electron. A similar expression may now be written for the angular acceleration of the negative elements, namely

$$\alpha' = \frac{2F'}{a'dm'}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

According to the mass-formula, the mass of a charge supposed spherical is inversely as its radius, the charge remaining the same. And, since each charge has been divided into the same number of elements, the masses of these elements are inversely as the radii of the said spheres, and again the radii of the spheres are directly as the radii of the two similarly placed elements.

Hence, for equal charges it is true that

$$\frac{a}{a'} = \frac{dm'}{dm}, \text{ or } a\,dm = a'dm'. \quad . \quad . \quad . \quad . \quad . \quad (9)$$

But for a positive charge of double the negative, since the radii vary as the squares of the charges, we have

$$a\,dm = 4a'dm'. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Hence the ratio of their angular accelerations by (7) and (8) is

$$\frac{\alpha}{\alpha'} = \frac{2F}{a\,dm} \times \frac{a'dm'}{2F'} = \frac{1}{4}\frac{F}{F'}. \quad . \quad . \quad . \quad . \quad (11)$$

But, by (6)

$$\frac{F}{F'} = \frac{d\mathrm{E}_1}{d\mathrm{E}_1'}\frac{q_1}{q_1'} = 2\frac{q_1}{q_1'} = 2\frac{\beta_1}{\beta_1'}. \quad . \quad . \quad . \quad . \quad (12)$$

Hence

$$\frac{\alpha}{\alpha'} = \frac{1}{2}\frac{\beta_1}{\beta_1'}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

But the gravitational equation, and the expression for the mass given in (68) of the first paper *, shows that the value of β for a ring equivalent to the positive charge of $2e$ is about 31·4 times the value of a corresponding ring equivalent to the negative electron. Hence the values of β for

* *Loc. cit.*

corresponding elements may be assumed to have this approximate ratio, and we have $\alpha = 15{\cdot}7\alpha'$ showing that the angular acceleration produced by the passing electron in the two elements of the positive charge is about 16 times greater than that produced in the corresponding elements of the negative electron. The total integral effect for all elements in each of these two bodies may be assumed to have approximately this same ratio, and hence the effect upon the negative electrons as above stated may be neglected and attention centred upon the positive charge.

The Effect of the Rotation of the Axis of the Positive Charge.

While it is true that this whole process cannot be worked out in exact detail because of the limitations of analysis, yet we are able to obtain a glimpse of what probably happens in a general way. If preferred, the processes now to be described may be regarded as assumptions, but with the understanding that they are clearly within the possibility of being substantiated by electrodynamics. Tracing the process as the passing electron has various velocities from zero up, it is clear that for a slow velocity the axis of C (fig. 1) may be turned slightly out of its normal position, the body of the positive charge rolling over smoothly on the surface of the electrons A and B without losing contact with them. Some energy must be taken from the passing electron if any alteration of this kind takes place in the atom. But whatever disturbance takes place upon the approach of the electron P, the normal condition is restored again as it departs, and all of the energy received is again returned to the electron, so that it passes the atom leaving with as much energy as it had upon approach.

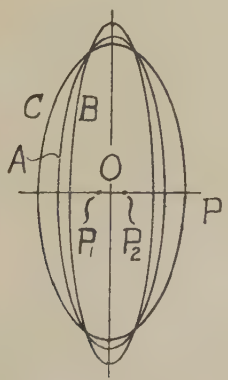
As the velocity of the passing electron becomes greater and greater the positive charge will be turned with more and more violence, and when the impulse is sufficient, it may turn with sufficient rapidity to leave contact with that electron, say A, whose pressure upon C is the least. This electron then returns and comes into actual collision with the positive charge, thereby setting up vibrations within its body that are characteristic of itself.

The case may be likened to rolling an elliptical wheel along a horizontal table. If the velocity of the wheel exceeds a definite critical value, the wheel will leave contact with the table, and upon its return bump against it, thus setting up sound waves. Some of the translational energy

of the wheel is then transformed into sound waves and is radiated out into space. So some of the translational energy of the passing electron is thus transformed in character into the high-frequency vibrations of the positive charge, and the energy thus transformed or received by the atom is not returnable to the passing electron because it has been thus transformed. This transformation and consequent reception of the energy by the atom is, therefore, only possible when the velocity of the passing electron exceeds a certain critical value, which prediction is in close accord with the known facts.

And further, just as is the case with the wheel set into vibration so as to emit waves of sound and thus radiate its energy into space to be reflected, refracted, and absorbed by all immediately surrounding objects, so the waves emitted by the vibrating positive charge are reflected from the objects surrounding it. The objects, however, immediately and practically completely surrounding it are the two negative electrons, A and B of fig. 1, and these effectively confine this received energy within the limits of the atom itself, none of it being radiated out into space because of these barriers.

Fig. 4.



Illustrating the change of shape when executing high-frequency vibrations natural to the positive charge. The effect of this vibration is taken as equivalent to two point charges, P_1 and P_2, vibrating from and toward each other on the axis OP to a first approximation.

Whatever the details of this complicated process may be, at a certain critical velocity of the passing electron the positive charge may be set into vibration, say like that assumed in fig. 4. The normal figure of the rotating positive charge is tentatively represented in section by the

ellipse A, its axis of rotation being OP. The vibration may be supposed to change its shape between the extreme limits of the curves B and C. Since the volume is supposed to remain constant, the surface area changes, and there occurs an interchange of energy or oscillation between the potential energy and the rotational energy, the potential energy being associated with the surface pressure, or the surface tension if preferred. The whole energy content of the positive charge thus takes part in this vibration.

The actual vibration of the positive charge is probably far more complicated than that represented in fig. 4, and it is possible that the fundamental vibration there represented does not take place at all. It must be true, however, that the value of the natural free frequency of the positive ellipsoid regarded as a solid body far transcends any that has ever been measured, for it has the free period of the smallest known body. It is not material to the arguments that follow what the precise value of this frequency is, provided it is very great; that is to say, much greater than the highest frequencies observed anywhere, and probably having a wave-length comparable with the dimensions of the negative electrons themselves. When this is the case the energy of vibration may be prevented from escaping from the atom, for it has been shown * that bodies having dimensions close to the order of magnitude of the received waves become great absorbers of this radiation. The present case differs from this, however, in that the word "absorbed" implies a transformation in the form of the energy, whereas the phenomenon above contemplated is more analogous to total reflexion, the whole energy being confined within the region between the two negative electrons of the atom and without any transformation in character.

The Effect of the Vibrating Positive Charge.

Let us next investigate the effect which a single positive charge of $2e$, supposed to be alone by itself and in vibration as above described, would have upon a negative electron in its immediate neighbourhood. It may be stated here in advance of the demonstration that these high-frequency vibrations give rise to a new ponderomotive force, which may easily exceed in value the great electrostatic forces that heretofore have controlled the situation, and that these forces act in such a direction as to annul the electrostatic forces and overpower them to such an extent that a negative

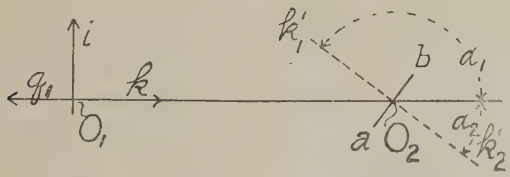
* L. Page, Astrophys. J. lii. p. 65, Sept. 1920.

electron may be driven away from its normal position in the atom.

To consider the positive charge as a vibrating oblate spheroid is beyond the power of analysis at the present time, but a first approximation may be made to the spheroid by substituting for it two vibrating points of charge of the same sign each equal to e. If the ellipsoid of fig. 4 is divided by the equatorial plane into two symmetrical halves and each half considered by itself, it is evident that the centre of volume of the charge of each half lies on the axis OP. These centres may be represented by the points P_1 and P_2. As the charge changes its shape during the vibration, these centres will move from and toward each other along the axis OP. It is thought that to a first approximation the whole effect of the vibrating charge may be obtained by replacing the volume charge by two point charges, P_1 and P_2, of the same sign and equal in value to half the positive charge.

This is not exactly the case of the electrical doublet upon which much work has been done because such doublets have been composed of charges of opposite sign, a positive and a negative charge. There has been no occasion heretofore, so far as is known, to treat of a doublet both parts of which have the same sign, and certainly such a case has never been treated by means of the Saha theory, which is used below. On this account it seems worth giving the demonstration quite fully.

Fig. 5.



Illustrating an electron at O_1 moving away from O_2 with velocity q_1, and a positive point charge vibrating with harmonic motion along *ab* about O_2 as centre.

In fig. 5 let O_1 be the position of a negative electron, and let O_2 represent the centre of motion of a positive point charge, $+e$, moving with an harmonic motion along the line *ab*, and between these points as limiting distances from the origin, O_2. It is required to find the average force upon the electron at O_1, which may be supposed to be moving directly away from O_2 with velocity q_1 due to the force

originating in the harmonically vibrating positive charge at O_2. Because of the small value of the rotation speed of the negative electron at O_1, we shall regard it merely as a point charge in translational motion, and may select three rectangular axes i, j, and k, such that the k-axis takes the direction O_1O_2, and the i-axis is in the plane determined by the line of vibration ab and the point O_1, that is the plane of the paper.

In this situation the fundamental ponderomotive force equation (1) tells us that the total force upon the negative electron at O_1 must lie in the plane of the paper and have no component along the axis of j, since the force is the sum of two vectors in the direction of $\mathbf{R}$ and of $\mathbf{q}_2$, both of which always lie in the plane of the paper, the i-k-plane. Since the problem of finding the force of a complete continuous ring of charge upon another has been worked out, we shall follow the general plan there used and make reference to the equations in the first paper *.

In order to solve the case of the simple harmonic motion let us begin with a circular motion and suppose that the harmonic motion of the point along ab of fig. 5 is replaced by a uniform circular motion around a circle of diameter ab, the plane of which is perpendicular to the paper. This merely adds a second harmonic motion along the axis of j, so that the motion of the second point may be expressed as before (*loc. cit.* equation (5))

$$\mathbf{r}_2 = a_2(S_2 i' + C_2 j). \quad . \quad . \quad . \quad . \quad . \quad (14)$$

The k'-axis is then perpendicular to the plane of this circular motion, that is in the plane of the paper, and the j and j' axes coincide as before. The angle, α, is that between the k and k' axes as shown in fig. 5. Using the formulæ (8) and (9) (*loc. cit.*) $\mathbf{r}_2$ is given by (10), namely

$$\mathbf{r}_2 = a_2[S_2(\cos\alpha)i + C_2 j + S_2(\sin\alpha)k]. \quad . \quad . \quad (15)$$

And by differentiation $\mathbf{q}_2$ is as given by (16) of the first paper,

$$\mathbf{q}_2 = a_2\omega_2[C_2(\cos\alpha)i - S_2 j + C_2(\sin\alpha)k]. \quad . \quad . \quad (16)$$

The velocity of the first point, O_1, is now, however, different, being motion along the k-axis, and may be represented by

$$\mathbf{q}_1 = -\beta_1 c k. \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

Hence

$$\frac{1}{c^2}\mathbf{q}_1 \cdot \mathbf{q}_2 = -\beta_1\beta_2 C_2 \sin\alpha_1. \quad . \quad . \quad . \quad (18)$$

(in place of (17) *loc. cit.*)

* *Loc. cit.*

The vector **R** now becomes instead of (11) (*loc. cit.*)

$$\mathbf{R} = -\mathbf{r} - \mathbf{r}_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

where

$$\mathbf{r} = zk \text{ (instead of (6) } loc.\ cit.) \quad . \quad . \quad . \quad (20)$$

Hence by (15) and (20)

$$\begin{aligned} \mathbf{R} = &-a_2 S_2(\cos\alpha)\, i \\ &-a_2 C_2\, j \\ &+ [-z - a_2 S_2(\sin\alpha)]\, k. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (21) \end{aligned}$$

Hence by (18) and (21)

$$\begin{aligned} (1 - \tfrac{1}{c^2}\mathbf{q}_1 \cdot \mathbf{q}_2)\mathbf{R} = &[-a_2 S_2 \cos\alpha - \beta_1\beta_2 a_2 S_2 C_2 \sin\alpha\cos\alpha]\, i \\ &+ [-a_2 C_2 - \beta_1\beta_2 a_2 C_2^2 \sin\alpha]\, j \\ + [-z - a_2 S_2 \sin\alpha - \beta_1\beta_2 C_2 z \sin\alpha - \beta_1\beta_2 a_2 S_2 C_2 \sin^2\alpha]\, k. & \quad (22) \end{aligned}$$

By definition

$$A_1 = 1 - \frac{\mathbf{q}_1 \cdot \mathbf{R}}{c\mathrm{R}}. \quad \text{(See (18) } loc.\ cit.)$$

Hence by (17) and (21)

$$\mathbf{q}_1 \cdot \mathbf{R} = \beta_1 cz + \beta_1 c a_2 S_2 \sin\alpha. \quad . \quad . \quad . \quad (23)$$

Whence

$$\frac{\mathrm{R}A_1}{c} = \frac{\mathrm{R}}{c} - \frac{\beta_1}{c}(z + a_2 S_2 \sin\alpha), \quad . \quad . \quad . \quad (24)$$

(instead of (19) *loc. cit.*)

The whole force upon the first point due to the second may now be found by substituting in (1) the expressions (22) and (24), giving

$$\begin{aligned} \mathbf{F} = \frac{(-e_1)(+e_2)(1-\beta_2^2)}{\mathrm{k}(\mathrm{R}A_2)^3(1-\beta_1^2)^{\frac{1}{2}}} \{ &[-a_2 S_2 \cos\alpha - \mathrm{R}\beta_2 C_2 \cos\alpha \\ &+ \beta_1\beta_2 z C_2 \cos\alpha]\, i \\ &+ [-a_2 C_2 - \beta_1\beta_2 a_2 \sin\alpha + \mathrm{R}\beta_2 S_2 - \beta_2\beta_2 z S_2]\, j \\ &+ [-z - a_2 S_2 \sin\alpha - \mathrm{R}\beta_2 C_2 \sin\alpha]\, k\} \\ & \quad . \quad . \quad . \quad (25) \end{aligned}$$

In this equation the distance a_2 is extremely small, being less than 10^{-16} cm., which is the order of magnitude of the positive charge, and we are only interested in the value of the force when R is large as compared with a_2, in those terms which vary as the inverse square of the distance, which are large as compared with higher orders in R.

Except for the R which appears indirectly in (25) through the sines and cosines, S_2 and C_2, it is correct to use the approximation that $R=s=r$. And, since O_2 is located on the k-axis in this particular case, $x=y=0$, and $r=z$. Therefore, omitting all terms not containing R or z in (25), as contributing nothing to the inverse square terms, and replacing x, y, and z by the direction-cosines X, Y, and Z, namely $X=Y=0$, and $Z=1$, (25) reduces to the following instead of (30) (*loc. cit.*):

$$\mathbf{F}=\frac{(-e_1)(+e_2)(1-\beta_2^2)r}{k(RA_2)^3(1-\beta_1^2)^{\frac{1}{2}}}\{[-\beta_2(1-\beta_1)C_2\cos\alpha]\,i + [\ \beta_2(1-\beta_1)S_2\ \ ]\,j + [\ \ -1-\beta_2C_2\sin\alpha]\,k\}, \quad (26)$$

where as in (31) (*loc. cit.*)

$$(RA_2)^{-3}=\left[1+3\beta_2\left(\frac{r}{r}\right)+6\beta_2^2\left(\frac{r}{r}\right)^2+10\beta_2^3\left(\frac{v}{r}\right)^3+\ldots\right]r^{-3}. \quad (27)$$

Corresponding to (21) and (22) (*loc. cit.*) the value of RA_2 becomes

$$RA_2=R-\frac{\mathbf{q}_2\cdot\mathbf{R}}{c}=R-\beta_2[-zC_2\sin\alpha]=R-\beta_2 v, \quad (28)$$

where

$$v=-zC_2\sin\alpha \quad (29)$$

and

$$\frac{v}{r}=-C_2\sin\alpha, \text{ for use in (27)}. \quad (30)$$

Equation (26) gives the whole force of the charge, $+e_2$, revolving in a circular orbit upon the negative electron, $-e_1$, moving along the axis of k. If the charge $+e_2$ is changed into an elementary charge dE_2, the expression becomes the equivalent of (42) (*loc. cit.*) * and may be integrated around the closed ring so as to give the whole force of the ring upon the retreating electron in a manner similar to that employed in the first paper. Although the distance r is now a function of time, whereas it was independent of time in the former example, yet this integration is made for one instant of time, and time is not the variable. The result of

* The r which appears in the denominator of (42) (*loc. cit.*) is a misprint. It should have been in the numerator of the coefficient as in (26) above.

the integration around the circle gives the force upon the electron due to the ring as follows:

$$\mathbf{F} = \frac{(-e_1)(+e_2)(1-\beta_2^2)}{kr^2(1-\beta_1^2)^{\frac{1}{2}}} \{ [\beta_2^2(1-\beta_1)\sin\alpha\cos\alpha + \tfrac{3}{2}\beta_2^4(1-\beta_1)\sin^3\alpha\cos\alpha]\, i + [-1-\tfrac{1}{2}\beta_2^2\sin^2\alpha - \tfrac{3}{8}\beta_2^4\sin^4\alpha \ldots.]\, k \}. \quad (31)$$

This corresponds to the final expression (57) in the first paper, which is the force of one ring upon another. By making $X=Y=0$, $Z=1$, and $\xi = X\cos\alpha + Z\sin\alpha = \sin\alpha$ in that equation we obtain (31) exactly except for the factor $(1-\beta_1)$ in the i-component. The presence of this β_1-term here is due to the new condition imposed, namely that the electron at O_1 in fig. 5 is in translational motion instead of being at rest relative to the centre O_2.

The force upon the electron at O_1 contributed by each and every element of this ring when averaged for time over a complete revolution or many revolutions must be the same as that due to any other element, provided the distance r does not change during one revolution. Consequently (31) expresses the average force upon the electron equally well whether the second charge is a single point revolving in a circle or whether it is a charge of the same amount uniformly distributed around the circle, provided the speed $\beta_1 c$ is zero. Moreover, it is clear that, if the distance r changes but little during a complete revolution of the point at O_2 as compared with its total value, then (31) may represent the average force upon the electron at O_1 even though it be in motion and β_1 is not zero. It may now be recalled that the frequency of revolution of the charge at O_2 is supposed in our example to be very great so that the wave-length of the radiated waves is comparable with the dimensions of the negative electron itself. This means that O_1 would only advance a distance comparable with its own dimensions even though it has a velocity as great as that of light, and this would make very little change in the distance r during a complete period. For this reason (31) may be taken to represent the average force upon O_1 with great approximation, even though its velocity is not small. But our problem is to find the effect of a simple harmonic motion of the point at O_2 upon the electron at O_1.

The circular motion of the charge at O_2 may be represented as the sum of two simple harmonic motions, the one

in the plane of the paper in fig. 5 and the other perpendicular to the paper. As above stated, the original equation (1) shows that the total force upon O_1 due to any kind of motion of O_2 confined to the plane of the paper, the i-k-plane, contributes no component of force upon O_1 in a direction perpendicular to this plane, the j-axis. Similarly it may be shown that any kind of motion of O_2 along the j-axis alone contributes a force upon O_1 along this same j-axis only due to such motion. The electrostatic part of the force is, of course, along the k-axis in every case. This may easily be shown because the term $\mathfrak{q}_1 . \mathfrak{q}_2$ in (1) vanishes when the two velocities are perpendicular to each other as now supposed, and the whole force, therefore, takes the direction of $\mathfrak{q}_2$ along the j-axis with the exception of the first or electrostatic term of (1).

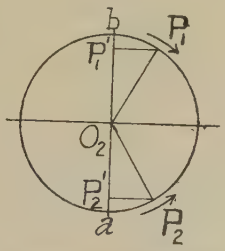
Now there is no j-component of the force in (31) above, which represents the result for circular motion, and since the component of this motion in the plane of the paper, fig. 5, can give no j-component, and the j-component is absent, it follows that the component harmonic motion along the j-axis contributes nothing at all to the total average force. The conclusion is that (31) represents equally well the average force upon the electron at O_1 whether the motion of the charge at O_2 is circular motion in a plane perpendicular to the paper in fig. 5 having a diameter ab, or is simple harmonic motion along the line ab. Hence (31) is the complete result for simple harmonic motion sought. The method employing the ring, which has been used above, seems to present less difficulty in effecting the integrations than would be the case were the time used as the independent variable, but the result must be the same.

The Effect of Two Vibrating Positive Charges.

In the problem just considered the vibrating charge at O_2 may be taken to represent, as above pointed out, one half only of the vibrating positive charge of the hydrogen atom. The other half vibrates in a similar manner with the same amplitude and frequency but in a different phase. The point P_1 in fig. 6 may represent the point above discussed revolving in the clockwise sense so that its projection P_1' describes the harmonic motion along ab. The other half of the positive charge may then be represented by a point, P_2, making the same angle with the diameter ab at all times and revolving in the counterclockwise sense, its projection being at P_2' describing harmonic motion along

the same diameter *ab*. To thus reverse the direction of rotation in the circle is equivalent to changing the direction of the axis of rotation through 180° in fig. 5 from the direction k_1' to that of k_2', thus changing the angle α_1 into α_2, the latter being the supplement of the former. Hence $\sin\alpha_1 = \sin\alpha_2$ and $\cos\alpha_1 = -\cos\alpha_2$.

Fig. 6.



Illustrating the opposite phases of the two points P_1 and P_2 of fig. 4.

The average force of the second vibrating point at O_2 in fig. 5 is, therefore, given by the same equation (31) by reversing the sign of the cosine but not that of the sine. This reverses the sign of the *i*-component but does not affect that of the *k*-component. In taking the sum of the forces due to both halves of the positive charge in harmonic motion, therefore, the *i*-component cancels and the *k*-component is double that expressed in (31). The total average force of the positive vibrating charge is then

$$F = \frac{2e^2(1-\beta_2^2)}{kr^2(1-\beta_1^2)^{\frac{1}{2}}}\left(1+\tfrac{1}{2}\beta_2^2\sin^2\alpha+\tfrac{3}{8}\beta_2^4\sin^4\alpha+\ldots.\right)k. \quad (32)$$

Since the series in the last parenthesis is the development of $(1-\beta_2^2\sin^2\alpha)^{-\frac{1}{2}}$, the complete force may also be written

$$F = \frac{e^2}{kr^2(1-\beta_1^2)^{\frac{1}{2}}}\left[\frac{2(1-\beta_2^2)}{(1-\beta_2^2\sin^2\alpha)^{\frac{1}{2}}}\right]k. \quad . \quad . \quad (33)$$

The Effect of the Hydrogen Ion upon an Electron.

The hydrogen ion is obtained by adding a negative electron to the positive charge above considered. By the fundamental equation (1) the force that this extra electron exerts upon the electron at O_1, fig. 5, is

$$F = -\frac{e^2}{kr^2(1-\beta_1^2)^{\frac{1}{2}}}k. \quad . \quad . \quad . \quad . \quad (34)$$

If this is added to (33) the force upon the electron due to the ion is as follows :

$$F=\frac{e^2}{kr^2(1-\beta_1^2)^{\frac{1}{2}}}\left[-1+\frac{2(1-\beta_2^2)}{(1-\beta_2^2\sin^2\alpha)^{\frac{1}{2}}}\right]k. \quad (35)$$

The combination of this ion with the negative electron at O_1 would make a neutral hydrogen atom, so that (35) represents the force upon the single electron which is being driven away from the atom in a straight line because of these new forces brought into existence by the high-frequency vibration of the positive charge.

Of course the analogy between this hypothetical case employing point charges and that where the electrons have finite dimensions must not be pressed too far, but it is believed that this approximate result will assist in understanding the process by which an electron may be driven out from its normal position in the atom to a maximum distance and back again in a straight line, thus radiating energy out into space and away from the atom. These equations seem worth examining in some detail on this account.

Let us first illustrate how the force depends upon the angle, α, shown in fig. 5, when other quantities are constant. When the forces in (33) for the positive charge and in (35) for the ion acting upon the negative electron at the distance, r, are plotted as polar curves with F and α as the only variables, a system of curves is obtained by giving to β_2 a succession of values, one curve for each value. It will be remembered that $\beta_2 c$ represents the maximum velocity of the harmonic motion of the vibrating points, P_1 and P_2, which have been taken as equivalent to the positive charge, and there is reason to think that this velocity may be very great initially, even close to the velocity of light, but that it decreases with time as the amplitude decreases with loss of energy because of the translational motion of the electron, and passes continuously through all values from its maximum down to its final velocity.

A system of six curves representing the force upon the electron due to the positive charge alone is shown in fig. 7 for the numerical values of β_2, 0, 0·5, $\sqrt{2}/2=0{\cdot}707$, 0·8, $\sqrt{3}/2=0{\cdot}866$ and 0·9 according to (33), which is the polar equation of an ellipse as regards the variables F and α. When $\beta_2=0$, the curve is a circle, since the force is merely the electrostatic force of attraction equal in all directions.

The radius of this circle is

$$2\frac{e^2}{\mathrm{k}r^2(1-\beta_1^2)^{\frac{1}{2}}}. \quad \ldots\ldots \quad (36)$$

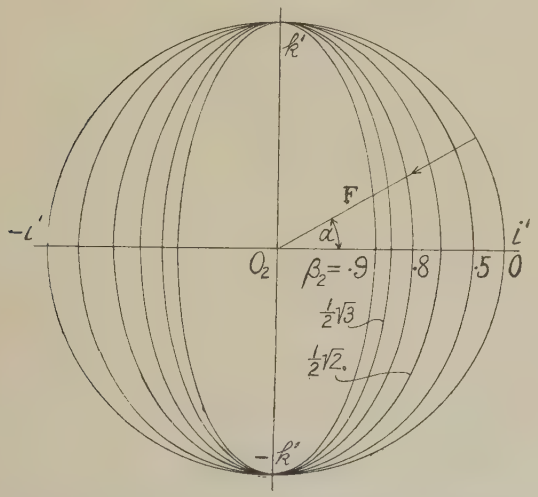
When $\beta_2=0{\cdot}5$, the major axis of the ellipse is

$$\sqrt{3}\frac{e^2}{\mathrm{k}r^2(1-\beta_1^2)^{\frac{1}{2}}}, \quad \ldots\ldots \quad (37)$$

and the minor axis

$$\frac{3}{2}\frac{e^2}{\mathrm{k}r^2(1-\beta_1^2)^{\frac{1}{2}}}. \quad \ldots\ldots \quad (38)$$

Fig. 7.

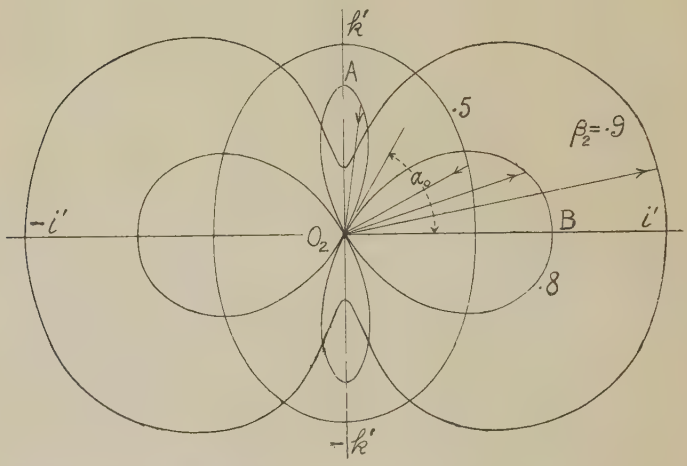


The radius vector from the curve to the origin O_2 represents the whole attractive force upon the electron at a fixed distance, r, due to a pair of positive point charges at O_2 vibrating in the direction O_2k'. Each ellipse shows the change in the force with the change in the angle α (see fig. 5), other quantities being constant. The six ellipses correspond to six different values of β_2 shown.

The major axes of all the ellipses are made equal to each other in the figure because the shape of the curve is the important feature rather than absolute values. The matter may be adjusted by supposing either that the scale of the drawing is changed in the above case in the ratio of $\sqrt{3}$ to 2, or that β_1 has different values in each of the curves. In the case when $\beta_2=0{\cdot}5$, $1-\beta_1^2$ may be changed in the ratio 4 to 3 in order that the major axis may have the same value as the radius of the circle for $\beta_2=0$.

The force due to the ion (35) in distinction to the positive charge alone is also represented by a circle when $\beta_2=0$, since it becomes the electrostatic force of half the value of (36). Three curves are shown in fig. 8 representing the force due to the ion corresponding to $\beta_2=0{\cdot}5$, $0{\cdot}8$, and $0{\cdot}9$ plotted from (35). They may be constructed

Fig. 8.



The radius vector represents the whole force upon an electron at a fixed distance, r, due to a pair of positive point charges at O_2 vibrating in the direction O_2k' together with one negative electron, the combination representing the hydrogen ion. Each curve shows the variation in the force with a change in the angle α (see fig. 5), other quantities being constant. When $\beta_2=\cdot5$ the force is attractive at all angles. At $\beta_2=\cdot8$ the force is attractive in the loop A and repulsive in the loop B. When $\beta_2=\cdot9$ the force is repulsive at all angles.

graphically from the corresponding ellipses of fig. 7 by subtracting a constant value from the radius vector of the ellipse equal in amount to the first term of (35). When $\beta_2=0{\cdot}5$, the force is still attractive at all angles α, but is greatest along the polar direction O_2k' and least along the equatorial direction O_2i'.

When $\beta_2=0{\cdot}8$, the force is still an attraction along the polar direction following the loop curve O_2A, but is a repulsion in the equatorial area within the loop O_2B. In this case there exists a critical angle, α_0, at which the whole force

vanishes. This angle is obtained by equating (35) to zero, giving

$$\frac{1-\beta_2^2}{(1-\beta_2^2\sin^2\alpha_0)^{\frac{1}{2}}}=\frac{1}{2}, \quad . \quad . \quad . \quad . \quad (39)$$

or

$$\sin^2\alpha_0=8-4\beta_2^2-\frac{3}{\beta_2^2}. \quad . \quad . \quad . \quad . \quad (40)$$

When $\beta_2=0{\cdot}8$, $\sin^2\alpha_0=0{\cdot}7525$, $\sin\alpha_0=0{\cdot}8675$ approx., and $\alpha_0=60°\ 10'$.

It is evident from the general expression in (39) that the force-curve will have two branches or loops, the one representing attractive forces nearer the pole and the other repulsive forces nearer the equator for a certain limited range of β_2, such that $\sin\alpha$ is real. The smallest possible value of β_2 to give any repulsive force may be found by making $\alpha=0$ in (39), giving $\beta_2=\sqrt{2}/2=0{\cdot}707$ approx. The curve obtained from (35) with this critical value of β_2 is shown in fig. 9 having two loops, O_2k', $-O_2k'$, showing a maximum attraction at the pole which falls to zero on the equator, a critical case.

As β_2 increases it is also seen by (39) that a second critical value is passed when $\beta_2=\sqrt{3}/2=0{\cdot}866$ approx., corresponding to $\sin\alpha=1$ and $\alpha=\pi/2$. The curve obtained with this value of β_2 is also shown in fig. 9 at O_2i', $-O_2i'$, representing two loops of repulsive force, the maximum being now on the equator, falling to zero at the pole, a second critical case.

As β_2 becomes larger than $\sqrt{3}/2$ the force is repulsive at all angles. At $\beta_2=0{\cdot}9$ the curve marked 0·9 in fig. 8 is obtained, showing a repulsive force even at the pole, though it is here a minimum. The scale of these different curves is not given as being unimportant, since it may be changed by varying the velocity $\beta_1 c$ of the negative electron.

Interpretation of the above Theory as applied to the Atom.

It appears at once from (35) that the repulsive force upon a negative electron produced by the vibration of the positive charge in the manner described may be large even when compared with the electrostatic forces which normally hold the electrons in close association with it. For, provided β_2 is great enough to render the bracket negative, that is in excess of $\sqrt{2}/2=0{\cdot}707$, the force is repulsion and its value depends upon the factor $(1-\beta_1^2)^{-\frac{1}{2}}$ in the coefficient, which

approaches infinity as the translational speed of the negative electron approaches the velocity of light. But the mass of the electron itself also approaches an infinite value in a similar manner and by the same form of function, for we have the well-known relation

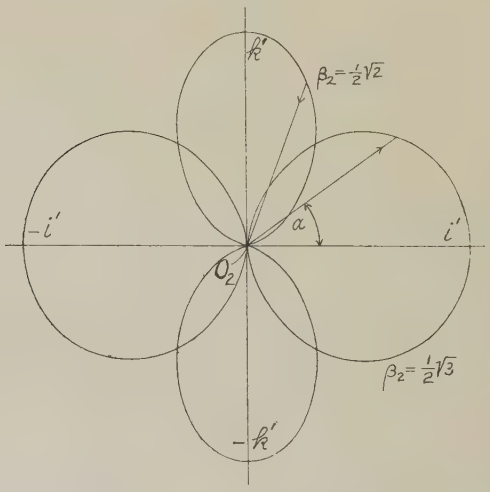
$$m = \frac{m_0}{(1-\beta_1^2)^{\frac{1}{2}}}, \quad \ldots \ldots \quad (41)$$

where m and m_0 are the masses at velocities $\beta_1 c$ and zero respectively.

Neglecting the reactive force due to the radiation from the electron in translational motion at present, and equating the whole force in (35) to the mass times the acceleration, the factor $(1-\beta_1^2)^{-\frac{1}{2}}$ cancels out, giving

$$\frac{d^2r}{t} = \frac{F}{m} = \frac{F(1-\beta_1^2)^{\frac{1}{2}}}{m_0} = \frac{e^2}{kr^2 m_0}\left[-1 + \frac{2(1-\beta_2^2)}{(1-\beta_2^2\sin^2\alpha)^{\frac{1}{2}}}\right]. \quad \ldots \quad (42)$$

Fig. 9.



Representing the same as fig. 8, but for two critical values of β_2, namely $\frac{1}{2}\sqrt{2}$ and $\frac{1}{2}\sqrt{3}$.

The above discussion, illustrated by the curves in figs. 7, 8, and 9, shows the important role that is played by the direction of the axis of the positive charge, represented in the formula by the angle α, for it is clear that this axis is free to be turned in response to the forces acting upon it

even during the brief interval when the electron has been temporarily separated from the positive charge. In the curve of fig. 8 corresponding to $\beta_2 = 0{\cdot}8$, for example, there is a critical angle α_0 at about $60^\circ\ 10'$ where the force changes sign from repulsion to attraction as the axis turns through this angle from a smaller to a greater value.

In this equation, (35) or (42), it is to be supposed that not only does the angle α vary with the time as the electron at O_1 (fig. 5) moves away from the atom, but that r, β_1 and β_2 are each functions of the time. It does not seem possible without further help to make valid assumptions as to these variables as functions of the time.

Some light upon the question is afforded by a consideration of the energy relations. Regarding the hydrogen ion at O_2 (fig. 5) as a single physical system, let us apply to it the ordinary rules of dynamics. Since the entropy of any material system left to itself is known to tend toward a maximum, this may be interpreted here to mean that the ion strives to be rid of its energy at a maximum rate. Now the kinetic energy of translational motion of the electron at O_1 is a part of the energy originally possessed by the positive charge of the ion. If the ion succeeds in driving this electron free completely away from itself carrying with it this kinetic energy, then its own energy will be reduced at a far greater rate than it would be if it depended merely upon the energy lost by radiation.

For, as will be shown later, the kinetic energy at some points of its path is much greater than the total energy radiated by the whole motion. To render the entropy a maximum, therefore, the ion strives to get rid of an electron. As shown above, there exists a certain critical angle α which reduces the whole force to zero, and if the ion could immediately turn itself to this angle, the force would remain zero and the translational velocity of the electron would then become constant because it would retain its original velocity of projection. No energy would be radiated out into space by this uniform velocity and zero acceleration, for the rate of radiation is zero for uniform motion, but the condition for maximum entropy would be fulfilled.

The legitimate inference from this reasoning is that the forces governing the direction of the axis of the positive charge strive to make it assume this critical angle, α_0, as soon as possible. But in so doing the mass of the ion, or its angular momentum, requires that a certain amount of time be consumed in this operation, causing the axis first to be carried beyond the critical angle and then to be

reversed in its motion, thus oscillating about the angle α_0 with a damped motion. The damping must be present because the value of β_2 is continually decreasing owing to loss of energy by radiation from the electron into space.

The maximum value of the periodic force must, according to fig. 8, be greater the greater the maximum departure from the critical angle, α_0, during a complete period. But the amplitude of this motion must be a function of its frequency, and it seems most logical to assume that the greater the frequency the smaller the amplitude of the motion. This may mean that the maximum force varies on this account alone as the inverse of the frequency.

The force according to (35) is also inversely as the square of the distance r, but r itself will be a periodic function of the time, and it is shown below that the maximum values of r vary inversely as the frequency. These two causes operating together, therefore, make the maximum values of the periodic force vary directly as the frequency of oscillation of the pole of the positive charge. This matter is stressed because it is unusual to find a force in mechanical systems whose maximum values vary as the frequency, but the above model points to this condition.

With the above understanding it seems entirely justifiable to assume that the acceleration expressed by (42) may be written as a simple periodic function of the time as follows :

$$\frac{d^2r}{dt^2} = k_1 \omega e^{-\omega t} \cos \omega t, \quad . \quad . \quad . \quad . \quad . \quad (43)$$

where ω is the angular velocity corresponding to a frequency, ν, such that $\omega = 2\pi\nu$. The presence of the factor ω in the coefficient makes the maximum values of the force, as above stated, directly proportional to the frequency ν, k_1 being a constant to be determined. The presence of a damping factor, $e^{-\omega t}$ has been accounted for by the loss of energy due to radiation from the electron moving with accelerated motion.

Granting that this is the typical form that the acceleration expressed by (42) would take, were it possible to deduce it as a function of the time alone, the velocity may be found by a first integration with respect to the time, giving

$$\frac{dr}{dt} = \tfrac{1}{2} k_1 e^{-\omega t} (\sin \omega t - \cos \omega t) + k_2, \quad . \quad . \quad . \quad (44)$$

k_2 being the constant of integration. When the time is infinite it may be assumed that the electron has returned to resume its original position in the atom, and that $dr/dt = 0$. Hence $k_2 = 0$, and no constant is required. A

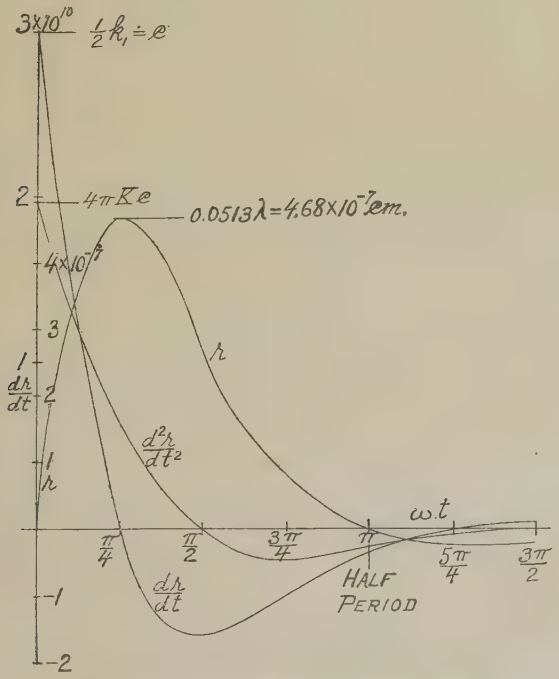
second integration gives the distance r, as follows:

$$r = -\frac{1}{2}\frac{k_1}{\omega} e^{-\omega t} \sin \omega t + k_3, \quad . \quad . \quad . \quad . \quad (45)$$

k_3 being a second constant of integration. This is also negligible because when the time is infinite $r_\infty = k_3$, and the distance from the centre of the atom to the normal centre of the electron in the atom is of the order of 10^{-13} cm. and quite negligible.

This expression for r shows that the maximum values vary inversely as the frequency other things being equal, which confirms the statement made above in regard to the distance r that occurs as r^{-2} in the expression for the force.

Fig. 10.



Three curves representing the acceleration, velocity, and distance respectively of the electron at O_1 (fig. 5), moving along the line joining it with the hydrogen ion at O_2 plotted as a function of the time. The maximum separation from the ion occurs at $\omega t = \pi/4$ and is equal to 0·0513 of the wave-length of the radiated wave. The total energy radiated is equal to one quantum, $h\nu$.

Three curves marked $\frac{d^2r}{dt^2}$, $\frac{dr}{dt}$, and r in fig. 10 represent the equations (43), (44), and (45) for certain assumed numerica

values of the constants. $k_1/2$ according to (44) is equal to the initial velocity of projection of the electron from the atom. The frequency ν is assumed for these curves to be equal to the Rydberg constant $3{\cdot}29 \times 10^{15}$ as representing the maximum frequency in the spectrum of hydrogen. The angles ωt are the abscissæ in each of the three curves.

The maximum distance of the electron from the ion is reached when $\omega t = \pi/4$, one eighth of a cycle, at which time the velocity is zero and reverses direction. Putting $\omega t = \pi/4$ in (45) and giving the initial velocity $k_1/2$ its maximum possible value, approximately the velocity of light, c, the maximum distance the electron is driven from the ion is given by

$$r_{\max.} = \frac{c}{\omega} e^{-\pi/4} \sin \frac{\pi}{4} = \frac{c}{\omega} \times 0{\cdot}456 \times 0{\cdot}707 = 0{\cdot}322 \frac{c}{\omega}. \quad (46)$$

Since $\omega = 2\pi\nu$, and $\lambda\nu = c$, we have

$$r_{\max.} = 0{\cdot}0513\,\lambda, \quad . \quad . \quad . \quad . \quad . \quad (47)$$

where λ is the wave-length of the radiated waves.

Further reference will be made to this result in a subsequent section.

The Total Energy Radiated during One Excursion of the Electron.

An expression for the rate of radiation of energy when a point charge moves in a curved path has been obtained by Liénard and has been shown by Schott * to agree with his deductions, both of these authors using the same form of electrodynamics, that of Larmor and Lorentz. Its expression is as follows :—

$$\dot{\mathrm{R}} = \tfrac{2}{3} c e^2 \left[\frac{\beta^4}{\rho^2 (1-\beta^2)^2} + \frac{\frac{1}{c^4}\left(\frac{d^2 r}{dt^2}\right)^2}{(1-\beta^2)^3} \right], \quad . \quad . \quad (48)$$

where ρ is the radius of curvature of the path, $\frac{d^2 r}{dt^2}$ the acceleration along the path, and β its velocity in terms of that of light, c.

The force on the electron caused by the reaction of the

* 'Electromagnetic Radiation.'

radiation in the direction of the tangent to its path follows from the definition of $\dot{\mathrm{R}}$, since

$$\dot{\mathrm{R}} = \frac{d\mathrm{E}}{dt} = \mathrm{F}\frac{dr}{dt} = \mathrm{F}v = \mathrm{F}\beta c, \quad . \quad . \quad . \quad (49)$$

with due regard to the sign according to whether the reaction of the radiation or a force equal and opposite to it is meant. Schott has given examples computing the total energy radiated when the motion is circular (*loc. cit.* p. 110), simple harmonic motion in a straight line (p. 122), and elliptical motion (p. 126), and shows that his results are obtainable from (48) above.

Since we have previously expressed and supported the view that the Larmor-Lorentz form of electrodynamics is in error in certain respects, and have shown that the Saha form of the theory removes these objections, it would be preferable now to use Saha's expression for the rate of radiation, were it known to the writer, instead of that of Liénard in (48). There seems scarcely any doubt that the form in (48) is defective or at the least incomplete when the case of simple harmonic motion in a straight line alone is considered. At the middle point of the motion the velocity is a maximum and the acceleration zero, hence at this point by (48) $\dot{\mathrm{R}}$ is zero as in uniform motion, and by (49), since dr/dt is not zero, F is zero, and there is no reaction due to radiation which is as it should be. As however, the point approaches the end of its harmonic path the velocity dr/dt approaches zero, and the acceleration approaches a maximum. Hence by (48) $\dot{\mathrm{R}}$ is a maximum toward the end of the path, but by (49) the reactive force equal to $\dot{\mathrm{R}}/v$ approaches infinity as close as we please, since v approaches zero as near as we please, while $\dot{\mathrm{R}}$ has its maximum value.

Physically this is unthinkable that the reactive force should approach so large a value in any part of the path, even though no difficulty is apparent when the total energy per cycle is computed. It is also realized that it is generally admitted that this formula only applies to an actual charge, an electron, which has a finite volume, provided the radius of curvature of its path is large as compared with the dimensions of the charge, but this restriction is hardly sufficient to remove the objections to the formula.

The fact that the expressions for the ponderomotive forces of a second moving point charge upon a first differ

so greatly as between the Lorentz and the Saha electrodynamics makes it quite certain that the Saha expression for the rate of radiation when obtained will differ materially from (48), but some interest may attach to the use of (48) in the meanwhile before the Saha expression is available. For straight line motion ρ is infinite and (48) becomes

$$\dot{R}=\frac{2e^2}{3kc^3}\frac{\left(\frac{d^2r}{dt^2}\right)^2}{(1-\beta^2)^3}. \quad . \quad . \quad . \quad . \quad . \quad (50)$$

The dielectric constant k is here inserted in order to keep the dimensions of both members of the equation the same. It should also have been written in (48), but that stands as quoted from the original author.

Assuming that the motion of the electron at O_1 (fig. 5) is completely described by the equations (43), (44), and (45) and illustrated by fig. 10, the total energy radiated during one whole excursion may be obtained from (50) by integration with respect to the time between the limits zero and infinity. Writing for β in (50) $\frac{1}{c}\frac{dr}{dt}$ we have the total energy

$$E=\int_0^\infty \dot{R}dt=\frac{2e^2}{3kc^3}\int_0^\infty\left[1-\frac{1}{c^2}\left(\frac{dr}{dt}\right)^2\right]^{-3}\left(\frac{d^2r}{dt^2}\right)^2 dt, \quad (51)$$

or, expanding into series, this becomes

$$E=\frac{2e^2}{3kc^3}\int_0^\infty\left[1+\frac{3}{c^2}\left(\frac{dr}{dt}\right)^2+\frac{6}{c^4}\left(\frac{dr}{dt}\right)^4 + \frac{10}{c^6}\left(\frac{dr}{dt}\right)^6+\ldots\right]\left(\frac{d^2r}{dt^2}\right)^2 dt. \quad . \quad . \quad (52)$$

Using the values of the acceleration and velocity given in (43) and (44) above and integrating each term separately gives finally for the total energy radiated in one excursion of the electron out and back

$$E=\frac{\pi e^2 k_1^2}{2kc^3}\left[1+\frac{23}{90}\left(\frac{k_1}{c}\right)^2+\ldots.\right]\nu. \quad . \quad . \quad . \quad . \quad (53)$$

The next term of this series will be k_1^4/c^4 with a different numerical coefficient, and so on, the powers of k_1/c increasing by 2 from term to term. But k_1 is constant by hypothesis according to (43), and there exist, therefore, no variables in the right member of (53) except the frequency of the

radiation, ν. Hence the whole energy may be written

$$E = h\nu, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (54)$$

which is Planck's quantum expression, if h may denote the coefficient of ν in (53). That is to say exactly one quantum of energy is radiated away from the atom by one excursion of one electron out and back. Moreover, the atomic model affords a concrete picture of just what it is that possesses this frequency, ν.

The numerical value of Planck's constant, h, cannot be deduced from (53) because the value of the constant k_1 is unknown, but the known value of h may be used to give some information about k_1. According to (44) it is seen that $k_1/2$ represents the initial velocity of ejection of the electron from the atom. The series in the parentheses of (53) may have any value from unity to infinity according to the value of k_1. It is evident by (50) that the radiated energy approaches an infinite value as the initial velocity of the electron approaches that of light. For all initial velocities less than that of light the series in (53) is finite, but less and less convergent as the velocity of light is approached. Suppose that the sum of the series within the parenthesis of (53) is denoted by S. In order that the coefficient of ν may be numerically equal to Planck's constant, h, then the value of S must be 138·2. This indicates immediately that a large number of terms of the series is required, that it is not very rapidly convergent, and that the initial velocity $k_1/2$ must be very close to that of light. Its exact value might be evaluated by the use of this information, but the process is laborious and not worth while. When $k_1 = 2c$ approximately, the constant h is then given by

$$h = 2\pi e^2 S/ck. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (55)$$

The Kinetic Energy of the Ejected Electron compared with the Total Energy Radiated.

It is seen by (44) that the maximum velocity of the electron upon its return toward the atom, after passing the point of maximum departure from it, is reached after one quarter of a period, when $\omega t = \pi/2$. Assuming an initial velocity close to that of light, its velocity at this point on return is about $0\,323 \times 10^{10}$ cm./sec., or not far from one tenth of the velocity of light. Its kinetic energy at this point is approximately 47×10^{-10} erg. The maximum energy radiated for the greatest light frequency that can

exist in hydrogen is $h\nu = h\mathrm{K} = 0{\cdot}2154 \times 10^{-10}$ erg, which is but the two hundred and twentieth part of the kinetic energy just computed, and this kinetic energy is much less than that at some other points.

The importance of the above conceptions, if true, is so great that some of the more fundamental points may bear restating. An initial velocity closely approaching that of light may be produced in an electron which is shot out in a straight line to a limited distance and then returns again to the atom. The energy thus lost by radiation from the atom is no greater than that received from the passing electron from which the energy was originally received. Because the form of this energy received from the passing electron is entirely transformed in character into high-frequency vibrations of the positive charge, the whole internal energy of the positive charge becomes available energy for the time being, and some of this great store of internal energy is used to give the ejected electron its large kinetic energy of translational motion, but this is all returned again to the atom. The positive charge undoubtedly contains a store of energy, $mc^2 = 1{\cdot}49 \times 10^{-3}$ erg, which is very great compared with any of the other energies mentioned. Hence it becomes possible for a very much smaller amount of energy, that of the passing electron, to release a much larger amount of energy, namely some of that stored within the positive charge. This may be called the trigger effect. Although such an effect has been mentioned in the literature as a remote possibility, I am not aware that its probability from the standpoint of electrodynamics has ever been pointed out. To admit of its existence will be shown in a subsequent communication to remove the usual difficulties connected with the photoelectric effect and secondary β-rays generated by X-rays. These are the two chief phenomena that seem to demand the artificial conception of "light quanta," which of itself introduces difficulties far more formidable than to admit of a trigger effect.

The Law of Force upon an Electron.

According to the above results the force acting upon the electron at precisely the same distance from the ion on its outward journey is very different from the force upon it at this distance on the return journey. The reason for this is clear, that the ion has changed its character in the meantime. The law of force according to this must not only be a function of the distance but of the time also. Attempts

have been made by various authors in the past to deduce a law of force as a function of the distance from experimental observations, but their results have been discordant. The laws found for one case do not work well for other cases and changed conditions. These discordant results are quite in agreement with the conception above mentioned that there is no such thing as a law of force dependent upon the distance only. The time must also enter into these functions.

The Absolute Values of the Maximum Distances to which an Electron is Projected from Hydrogen.

According to (47) the maximum distance to which an electron goes from the atom is proportional to the wave-length of the radiated waves, λ. The shortest wave-length of the hydrogen spectrum is $911{\cdot}78 \times 10^{-8}$ cm., corresponding to the frequency $K = 3{\cdot}28880 \times 10^{15}$, the Rydberg constant. Hence in this case the distance given by (47) is $4{\cdot}677 \times 10^{-7}$ cm. If the hydrogen gas in the vacuum tube is at 0° C. and the pressure is one millimetre of mercury, the approximate distance between the molecules of the gas on the average is about $65{\cdot}5 \times 10^{-7}$ cm., or about fourteen times greater than the maximum distance computed above. On the average, therefore, the ejected electron will not come within the sphere of influence of the next molecule of the gas, and will experience no disturbance or interference from this cause.

As the frequency of the characteristic radiation is reduced and the wave-length increased the maximum distance becomes greater. The wave-length must be increased fourteen times in order that the electron may go on the average as far as the next molecule. This would make the wave-length about $12{,}760 \times 10^{-8}$ cm., which is not far from the second line of the Paschen series in the infra-red, so that practically the whole of the observed spectrum of hydrogen may be emitted before an electron travels so far as the next adjacent molecule on the average, and no interference will thus be experienced.

Concluding Remarks.

The complex nature of the behaviour of the atom contemplated in this paper has seemed to require a presentation less abbreviated than is customary in order that the meaning intended may be made clear. The phenomena that have been described are of a fundamental nature, and have a

wide application in every phase of atomic action. It is believed that the conceptions of a high-frequency vibration of the elementary charges themselves, regarded as quasi-solid bodies, and of radiated energy taking place in quanta because an electron is temporarily shot out from an atom and in most cases returns to it again, will prove fruitful in clearing up many phenomena now obscure.

The radiation emanating from these ejected electrons, which falls upon an atom in the eye or in a photographic plate exposed to it, can be shown by means of the ponderomotive force equation (1) to exert a force upon the positive charge of such atoms having the same periodicity as that belonging to the ejected electrons by virtue of their accelerated motion (43). It is almost axiomatic, not requiring any detailed proof, that this frequency or period is transmitted through without alteration and makes itself felt upon the receiver.

The question as to how the line spectrum of hydrogen or any other element may be produced has not been treated because it would extend the paper to undue length, but the derivation of the frequency equation of the line spectrum of hydrogen may be obtained in a manner analogous to that suggested by Whittaker * in 1911. He then attributed the spectrum to the gyroscopic connexion between the two atoms of the hydrogen molecule, and deduced the equation of the Balmer series on this basis. Instead of the gyroscopic connexion being between the two atoms of the molecule it is between the electron ejected from the atom and the remainder of the atom itself, that is the ion, for it is now quite certain that the atom rather than the molecule is the source of these spectra.

In closing it may be emphasized that the treatment in this and the three preceding papers points to the probability that the total energy content of the elementary positive and negative charges is of the utmost importance, as being the key to most atomic actions, because it is drawn upon temporarily in radiation, in the dissociation of the atoms of a molecule, that is all chemical reactions, and probably in all phenomena. Its value, mc^2, is so great by comparison with all observed cases of energy transfer that temporary changes in it escape observation.

In a subsequent communication it is hoped to show how the ideas above presented throw light upon other phenomena not mentioned here.

* E. T. Whittaker, Proc. Roy. Soc. ser. A, vol. lxxxv. No. A 578, p. 262, June 9, 1911.

III. *On the Theory of Ionization by Swiftly Moving Electrified Particles and the Production of Characteristic X-Rays.* *By* Svein Rosseland *.

Introduction.

A COLLISION between an electrified particle and an atom will in general not conform to the laws of classical dynamics. The deviations from these laws become especially apparent when considering electrons moving with small velocities in inert gases or metallic vapours. In this case, as was first shown by Franck and Hertz, the electron will lose no energy at all if the velocity is below a certain critical value, while there is a considerable probability that it will lose all its energy in a single collision if this value is exceeded. As is well known, this effect is in accordance with the quantum theory of atomic structure. In fact, when a slowly moving electron collides with an atom in the normal state, we may on the basis of Bohr's theory expect that if the electron loses energy in the collision this energy must correspond to the transference of the atom to some other stationary state, and in experiments it is found that the critical velocity just corresponds to the transference of the atom from the normal to the next stationary state †. On the other hand, if the velocity of the incident particle is large compared with the velocities of the electrons in the atom, it is to be expected that the effect of a collision may be calculated with great approximation on the basis of the classical system of mechanics. A strong support for this view is, among other things, provided by the theory of ionization produced by swiftly moving electrified particles. In this paper we shall consider some aspects of this theory and its bearing on the production of characteristic X-rays.

I. Ionization by impact of α and β particles.

§1. *General Theory.*

A theory of the ionization produced by α and β rays was first proposed by J. J. Thomson ‡. In this theory the electrons are considered as being initially at rest and the mutual action between the incident particle and each of

* Communicated by Prof. N. Bohr.

† Compare J. Franck and G. Hertz, *Physikalische Zeitschrift*, xx. p. 132 (1919).

‡ J. J. Thomson, Phil. Mag. xxviii. p. 449 (1912).

the electrons is calculated as if the electron were uninfluenced by interatomic forces, supposing the interaction between the particle and an electron to conform to the law of the inverse square. It is easily calculated that the energy Q transferred to an electron in a collision of this kind is

$$Q=\frac{2m_1m_2^2}{(m_1+m_2)^2}V^2\sin^2\frac{\theta}{2}\,;\quad \tan\frac{\theta}{2}=\frac{eE}{V^2}\frac{m_1+m_2}{pm_1m_2}. \quad . \quad (1)$$

where m_1, $-e$, m_2, E are the mass and charge of the electron and the particle respectively, while p is the least distance from the initial position of the electron to the path of the particle before the collision ; θ is the angle of deflexion of the particle in its motion relative to the common centre of gravity. Putting $Q=W$ the energy necessary to remove the electron from the atom and denoting the corresponding p by p_W, we obtain

$$p_W^2=\frac{2e^2E^2}{m_1V^2}\left(\frac{1}{W}-\frac{(m_1+m_2)^2}{2m_1m_2^2V^2}\right). \quad . \quad . \quad . \quad (2)$$

In the following p_W will be termed "the radius of ionization." The number of collisions per unit path experienced by the particle with the electrons of a particular group of the atom, in which the energy transferred to the electron is larger than W, is πp_W^2Nn, where N is the number of atoms per unit volume and n the number of electrons in the group. Thus the total number of electrons removed from the group by the incident particle per unit path is

$$I=\frac{2\pi e^2E^2Nn}{m_1V^2}\left(\frac{1}{W}-\frac{(m_1+m_2)^2}{2m_1m_2^2V^2}\right). \quad . \quad . \quad . \quad (3)$$

This expression was by Thomson taken to be equal to the number of pairs of ions produced by the particle per unit path. As observed by Thomson, this formula gives a close representation of the relative number of ions produced by α and β particles, as well as the dependence of the ionization upon the velocity of the particle.

The total number of pairs of ions produced by α rays from RaC in air per unit path was determined by Geiger *. From the measurements of Taylor on the relative ionization produced by α rays in air and hydrogen, Bohr † calculated that the absolute number of ions produced by an α particle per cm. at the initial part of its range was several times larger than that predicted by (3). This excess effect was

* H. Geiger, Proc. Roy. Soc. A, lxxxii. p. 486 (1909).
† N. Bohr, Phil. Mag. xxx. p. 606 (1915).

attributed to the secondary ionization produced by the electrons expelled from the atoms. As the velocity of these electrons will be considerably smaller than that of the primary particles, the ionization produced by these electrons cannot be calculated on the simple assumption underlying formula (3), as the problem in reality is one of quantum dynamics. Using some simplifying assumptions Bohr deduced a formula giving an upper limit for the secondary ionization to be expected, and showed that this formula predicted the right order of magnitude of the ionization produced by α rays in hydrogen and helium. This result is of great interest, as it indicates, as presupposed, that for high velocities of the incident particles, and for violent collisions resulting in the removal of an electron from the atom, calculations based on the classical dynamical laws will give also the absolute value of the primary ionization.

When the velocity of the particle becomes of the same order of magnitude as that of some electrons in the atom with which it collides, formula (3) loses all claim of validity, inasmuch as the assumption that the electrons are initially at rest is not even approximately fulfilled, while at the same time the experiments of Franck and Hertz show that it is precisely in this region that the peculiar deviations from the laws of ordinary mechanics become apparent. It will also be found that the formula of Thomson does not even approximately correspond to experiment in this region; but it is after all interesting to note that it predicts a qualitatively right result, as is seen upon comparing it with the measurements of Kossel * and Mayer † of the ionization produced in gases by slowly moving electrons, as well as by α rays. The point is that both theory as well as experiment demand that the ionization shall increase initially to a sharp maximum and next decrease more slowly as the velocity of the particles is further increased. The exact position of the maximum is extremely difficult to determine experimentally, but it is found to occur when the velocity of the incident particles is a few times larger than that of the electrons in the atom. The ionization curve given by Thomson's formula rises much more rapidly to a maximal value than the experimental curves. Considering the experimental difficulties associated with the determination of the elementary ionization on this part of the curve, this disagreement is probably not larger than might be expected already on the

* W. Kossel, *Ann. d. Phys.* xxxvii. p. 393 (1912).
† F. Mayer, *Ann. d. Phys.* xlv. p. 1 (1914).

basis of the classical system of mechanics, when we remember that the electrons in the atom are revolving with velocities which on this part of the curve are of the same order of magnitude as the velocity of the incident particles.

§ 2. *Multivalent ions produced by a single impact.*

Consider a particle penetrating through an atom of some substance. Around the path of the particle as an axis, supposed to be a straight line, imagine a cylindrical surface with a radius equal to the radius of ionization of a definite electronic group. Suppose, for simplicity, the electrons of the group independently distributed on the surface of a spherical shell with radius r. Let σ denote the ratio of the part of the spherical surface situated within the cylinder to the area of the whole sphere.

The probability that out of the n electrons of the group a number m shall at a time be situated within σ is then

$$P=\frac{n!\,\sigma^m}{m!\,(n-m)!}(1-\sigma)^{n-m}. \quad . \quad . \quad . \quad . \quad (4)$$

Let q be the distance from the centre of the group to the axis of the cylinder. The mean number of collisions per unit path in which q lies between q and $q+dq$ is $2\pi Nqdq$, where N is the number of atoms per cm. The number of m multiple ions produced is accordingly

$$I_m=\frac{2\pi Nn!}{m!(n-m)!}\int_0^{r+p_W}\sigma^m(1-\sigma)^{n-m}qdq. \quad . \quad . \quad (5)$$

Let us fix our attention on the case in which the electronic group contains two electrons, corresponding to the neutral helium atom or the K-group of the other elements. The exact value of I_m even in this case is seen to be very complicated. We shall therefore be satisfied with approximate expressions. When p is small compared with r, the ratio σ will be small and we may approximately put

$$\sigma=\frac{1}{4\pi r^2}\frac{2\pi {p_W}^2}{\sqrt{1-\frac{q^2}{r^2}}}=\frac{{p_W}^2}{2r^2}\frac{1}{\sqrt{1-\frac{q^2}{r^2}}}, \quad . \quad . \quad (6)$$

and take the limits of the integral to be 0 and $r-p_W$. The

number of doubles is

$$I_2 = -\frac{\pi}{4} N \frac{p_2^4}{r^2} \log_n \left(\frac{2p_2}{r} - \frac{p_2^2}{r^2} \right) \quad . \quad . \quad . \quad (7)$$

and of singles

$$I_1 = 2\pi N p_1^2 \left\{ 1 + \frac{p_1^2}{4r^2} \log_n \left(\frac{2p_1}{r} - \frac{p_1^2}{r^2} \right) \right\} . \quad . \quad (8)$$

From the experiments of Franck and Knipping * it is known that the ionization potential of helium, when both electrons are removed at one time, is more than double of the ionization potential of one electron. The values of p in the formulæ (9) and (10) are therefore marked with different indices. The percentage of doubles is approximately

$$\epsilon = \frac{1}{8} \frac{p_2^4}{p_1^2 r^2} \log_n \left(\frac{2p_2}{r} - \frac{p_2^2}{r^2} \right) . \quad . \quad . \quad (9)$$

When, on the other hand, p_2 is larger than r we may put $\sigma = 1$ for $q < p - r$ and $\sigma = \frac{1}{2}\left(1 + \frac{p-q}{r}\right)$ when $p - r < q < p + r$. The number of doubles, singles, and percentage of doubles are found by an easy calculation to be

$$I_2 = \pi N [p_2^2 + \tfrac{1}{3} r^2 - \tfrac{5}{6} p_2 r] \; ; \quad I_1 = \tfrac{4}{3} \pi N p_1 r ;$$

$$\epsilon = \frac{p_2^2 + \frac{1}{3} r^2 - \frac{5}{6} p_2 r}{p_2^2 + \frac{1}{3} r^2 + (\frac{4}{3} p_1 - \frac{5}{6} p_2) r} . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

In the region where p and r are of the same order of magnitude neither (9) nor (10) will apply. However, drawing a diagram by means of these formulæ for large and small values of p/r, the probable course of the curve for moderate values of p/r will be given within narrow limits by an interpolation as shown in fig. 1, used in the next section.

§ 3. *Dimensions of the Helium atom.*

The preceding calculations are capable of direct application to the experiments of Millikan † and Wilkins ‡ on the production of divalent ions in helium. It was found that the percentage of divalent ions varied roughly in the same manner with range as the total ionization in the ordinary Bragg curve, and increased from about 8 per cent. at the

* J. Franck and P. Knipping, *Phys. Zeitschr.* xx. p. 7 (1919).
† R. Millikan, Phys. Rev. xviii. p. 456 (1921).
‡ T. R. Wilkins, Phys. Rev. xix. p. 210 (1922).

middle of the range of the alpha particle to about 16 per cent. at the top of the ionization curve, after which the percentage decreased rapidly to zero. As the only unknown quantity in the above expressions is r, these experiments will yield a value for the radius of the helium atom.

In order to calculate the radius r from the experiments we must first calculate the values of p_1 and p_2. Limiting ourselves to a region where Thomson's formula may be considered to be valid, we first notice that the ratio p_1^2/p_2^2 is nearly equal to the ratio P_2/P_1, where P_1, P_2 are the ionization potentials for one electron of helium, according as one electron or both are removed from the atom in one collision. From the work of F. Horton and A. C. Davies * and J. Franck and Knipping † it is known that P_1 is close to 25·4 volts. When this value is taken for granted the value of P_2 is unambiguously given by Bohr's theory as $\frac{1}{3}(4.13{\cdot}5 + 25{\cdot}4) = 39{\cdot}8$ volts. In fact, the experimental value given by Franck and Knipping is in complete accord with this value.

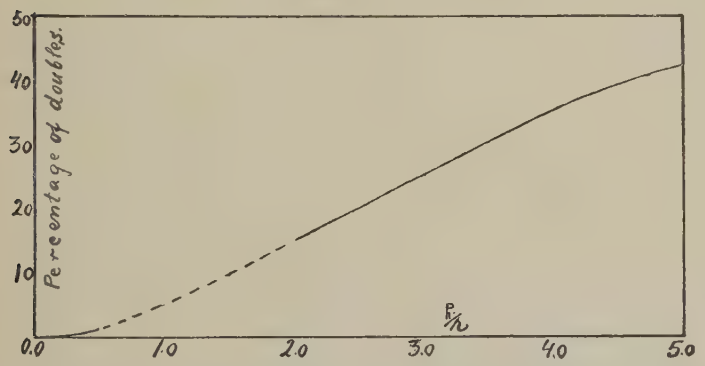
We now calculate the percentage of double catches for small and large values of p_1/r by means of the formulæ (9) and (10) and interpolate for intermediate values. Knowing the radius r, the absolute values of the percentage may be obtained by introducing the values of p_1 and p_2 calculated from Thomson's formula at a part of the range where this formula may be assumed to be valid. However, we shall prefer by reversing the argument to calculate the radius r and compare the result with the values given by the kinetic theory of gases and the quantum theory.

The velocity of the alpha particles from polonium is $V = 1{\cdot}55.10^9$ cm. sec.$^{-1}$. At the middle of the range it has decreased to $V = 1{\cdot}25.10^9$ cm. sec.$^{-1}$. From the observed value of the ionization potential we then calculate the corresponding p_1 to be $2{\cdot}93.10^{-9}$ cm. Millikan ‡ states that at the middle of the range the percentage of doubles is about 8 per cent. From the curve (fig. 1) the corresponding ratio p_1/r is seen to be close to 1·25. By means of the calculated value of p_1, this gives $r = 2{\cdot}5.10^{-9}$ cm. This value is four times smaller than the value afforded by the kinetic theory of gases (ca. 10^{-8} cm.) §, but it will be seen to be of the same order of magnitude as that predicted by Bohr's theory. If the two electrons revolve in circular one-quantum orbits,

* F. Horton and A. C. Davies, Proc. Roy. Soc. xcv. p. 408 (1919).
† J. Franck and Knipping, *Phys. Zs.* xx. p. 5 (1919).
‡ R. Millikan, Phys. Rev. xviii. p. 458 (1921)
§ Jeans, ' Dynamical Theory of Gases,' sec. ed. p. 341.

the radius of the orbits will on this theory be close to 3.10^{-9} cm.* This figure will be only little affected, whether we assume the orbits to be coplanar or inclined to each other. Conversely, assuming the value of r to be that given by the quantum theory, the percentage of doubles at the middle of the range comes out about 5 per cent. This value differs from that given by Millikan by scarcely more than the experimental error. These calculations thus seem strongly to suggest that the dimensions of the electronic orbits in the helium atom really are as small as demanded by Bohr's theory. Experiments of this kind seem to be of special interest, since we are here more directly concerned with the

Fig. 1.



dimensions of the electronic orbits, and need not take account of the complicated phenomena associated with a collision between two neutral atoms.

Our calculations are of course to some extent affected by the special distribution of electrons assumed. For the most unfavourable case when the two electrons constantly are placed at the opposite ends of a diameter through the nucleus, the probability of double ionization is easily seen to be of the order

$$\epsilon = \frac{1}{16}\frac{{p_2}^4}{{p_1}^2 r^2}, \qquad \ldots\ldots (11)$$

if p_2 is not larger than r. For the present case the percentage of doubles predicted by this expression will be about half as large as that predicted by the combination of (9) and (10). Although this result at present hardly permits of drawing any definite conclusions, it indicates that accurate

* N. Bohr, Phil. Mag. xxvi. p. 476 (1913).

experiments of this kind may be of value for an investigation of the configuration of the electrons within the group of the helium atom. The possible bearing of these experiments on the electronic configurations in helium was pointed out by Millikan *, but it is to be emphasized that his arguments are not conclusive without fixing the value of the atomic dimensions, and especially that no conclusion of this kind can so far be drawn from the observations near the optimal ionization. In fact, in this place the velocity of the α particles has become of the same order of magnitude as the velocity of the electrons in the atom. The motions of the two electrons during a collision cannot then be treated separately, and it is not certain that differences in the initial motion of the two electrons of the kind considered by Millikan will have any sensible influence on the probability of double catches.

II. Intensities of characteristic X-rays.

§ 1. *Production of X-ray spectra on the quantum theory.*

According to Bohr's theory of atomic structure † the electrons in an atom in the normal state are divided into groups in such a way that the orbits of the electrons in each group are characterized by a principal quantum number, the value of which is equal to one for the innermost group and increases by one unit when we pass from group to group in the direction from the nucleus outwards.

When a high speed electron is projected through the atom electrons may be ejected from the different groups. The rest of the atom reorganizes provisionally into a stationary state, in which, however, the number of electrons in the different groups is now different from that corresponding to the most stable arrangement of the remaining electrons. The atom reorganizes ultimately into a final stable state through a series of successive processes of transition, during each of which an electron is transferred from an outer to an inner group characterized by a smaller principal quantum number. Each of these processes is accompanied by the emission of a train of harmonic waves of frequency ν given by the relation $h\nu = E' - E''$, where E' and E'' are the values of the energy of the atom in the initial and the final states.

The intensities of the characteristic X-rays will primarily depend upon two circumstances. Firstly upon the number

* R. Millikan, Phys. Rev. xviii. p. 459 (1921).

† See for instance N. Bohr, *Zs. für Phys.* ix. p. 1 (1922).

N_i of atoms in the anticathode in a particular stationary state i prepared for X-ray emission ; and secondly upon the probabilities A_k^i of spontaneous transition * from the given stationary state i with a given distribution of the electrons in the different groups, to another stationary state k with the same total number of electrons in the atom, but where one electron has been transferred from an outer to an inner group. If ν_{ik} is the frequency of the radiation emitted by a transition between the stationary states i and k, the energy emitted in the form of radiation of this frequency in the time interval dt from the atoms of the anticathode is

$$dI = N_i A_k^i h\nu_{ik} dt. \quad . \quad . \quad . \quad . \quad . \quad (12)$$

In this expression the A_k^i are constants of the atomic configurations, which, on account of the fundamental assumptions underlying the quantum theory in its application to atomic structure, are necessarily independent of the manner in which the stationary state in question is brought about, *i. e.*, in the case of the excitation of characteristic X-rays, independent of the voltage applied to the X-ray tube.

Whatever function N_i may be of the voltage applied to the X-ray tube, the form of expression (14) involves an immediate conclusion concerning the relative intensities of all X-ray lines emitted when atoms in a specified stationary state i suffer spontaneous transitions into a series of stationary states k. For since N_i is the same for all these lines, the relative intensities of these lines will depend only upon the different products $A_k^i \cdot \nu_{ik}$, and thus will be independent of the voltage applied to the tube. This conclusion is generally confirmed by experiment.

In the general case N_i signifies the number of atoms with an arbitrary number of electrons absent from the different electronic groups. From the actually observed X-ray spectra it is seen that the majority of the lines are emitted from atoms that have lost but a single electron from the inner groups. However, as pointed out by Wentzel †, a number of satellites appearing in the lightest elements, which hitherto have escaped classification, may be explained by assuming that they are emitted by atoms that have lost two or more electrons from the inner groups. But since the intensities of these lines are small compared with that of the principal lines of the spectrum, and as they preferentially appear in the spectra of light elements, we shall in the

* A. Einstein, *Physikalische Zeitschrift*, 1917, p. 122.
† G. Wentzel, *Ann. d. Phys.* lxvi. p. 437 (1921).

discussion of the dependence of intensity of the characteristic X-rays upon voltage provisionally neglect the distinction between single and multiple ionization.

§ 2. *Dependence of intensity of X-ray lines upon the voltage applied to the X-ray tube.*

Consider an electron moving an element of path dx in the anticathode. During the motion it will collide with the atoms lying close to the path and lose energy to the electrons of the atoms. By far the greater part of the loss of energy will be due to collisions with the numerous loosely bound electrons belonging to the outer groups of the atoms. But there will also be a number of collisions with the strongly bound electrons of the inner groups. Let I(E) be the number of collisions with electrons of some inner group experienced per unit path by an electron moving with a kinetic energy E, in which these electrons are actually ejected from the atom. The total number of electrons removed from the group in question is then obtained by integrating $I\,dx$ along the path of the particle until the energy of the particle is just equal to the energy of ionization of the group in question. The mean decrease of energy per unit path of a swiftly moving electron is known from theory as well as from experiments to be of the form *

$$\frac{dE}{dx} = k \cdot \frac{N}{\frac{m}{2} \cdot V^2}, \quad \ldots\ldots \quad (13)$$

where k is a function of E increasing slowly with decreasing E, but which for the present purpose may be treated as a constant. V is the velocity of the electron and E is for all velocities given by the relativity formula

$$E = mc^2 \left(\sqrt{\frac{1}{1 - \frac{V^2}{c^2}}} - 1 \right)$$

where c is the velocity of light.

For voltages of the same order of magnitude as W the path travelled in the anticathode will be so small that it is unlikely that the absorption of the X-rays in the anticathode will influence the intensity of the emitted radiation to any appreciable extent. Similarly, the scattering of the electrons within the anticathode is immaterial in this connexion as

* See the discussion by Bohr, Phil. Mag. xxx. p. 594 (1915).

long as we can neglect the effect of the absorption of the emitted radiation in the anticathode. From experiments, however, the scattering is known to be so large that an appreciable number of electrons is actually reflected from the target. The velocities of these electrons have been measured and found to be essentially smaller than that of the incident electrons. For our purpose all electrons moving with a kinetic energy smaller than the ionization energy W_i are of no consequence. This circumstance is of essential importance, and we must expect to obtain results of the right order of magnitude when neglecting the fraction of the incident particles reflected with an energy greater than W_i. The number of electrons thrown out from the electronic group considered per unit time is

$$z\int I(E)dx. \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

where ze is the current passing through the tube. Introducing E as an independent variable from (12), we obtain

$$N_i = \frac{z}{kN\Sigma_s A_s{}^i}\int_{W_i}^{E_0} I(E)\tfrac{1}{2}mV^2 dE. \quad . \quad (15)$$

where $\Sigma_s A_s{}^i$ is the total probability of spontaneous transition from the state i. From (15) we obtain

$$I(E) = \frac{kN\Sigma_s A_s{}^i}{\tfrac{1}{2}mV^2} z\frac{\partial N_i}{\partial E}. \quad . \quad . \quad . \quad . \quad (16)$$

The intensity of a certain X-ray line corresponding to the state of ionization characterized with the letter i is

$$F_i = A_k{}^i h\nu_{ik} N_i, \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

We may therefore write

$$I(E) = \frac{kN}{a_{ik} z h\nu_{ik}} \frac{1}{\tfrac{1}{2}mV^2} \frac{\partial F_i}{\partial E}. \quad . \quad . \quad (18)$$

where $a_{ik} = \dfrac{A_k{}^i}{\Sigma_s A^{si}}$ is the relative probability of spontaneous transition from the state i to the state k.

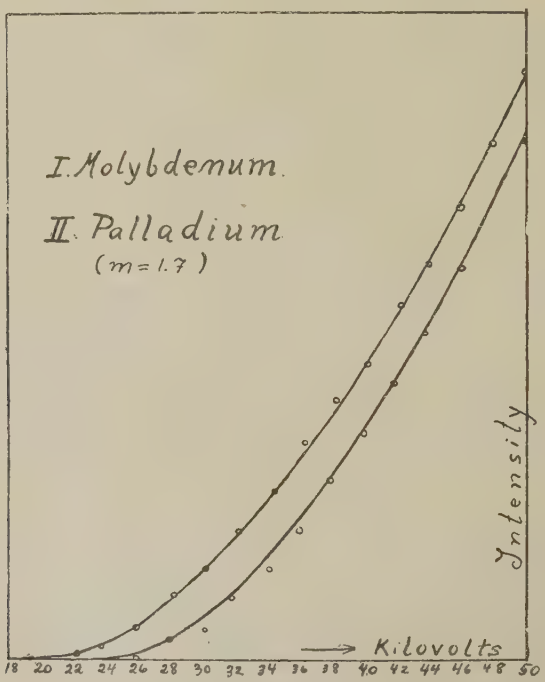
By means of this equation we can deduce I(E) from the experimental voltage-intensity curves, as long as E is not so large compared with W_i that the absorption in the anticathode of the excited radiation must be taken into account.

Experimental voltage-intensity curves have been obtained by D. Webster * for the K_α and K_β lines of rhodium, by

* D. Webster, Phys. Rev. vii. p. 599 (1916).

Webster and Clark * in the case of the L_α-lines of platinum, and by B. A. Wooten † in the case of the K_α and K_β lines of molybdenum and palladium. These curves can be represented by functions of the form $F = \text{const.}\,(E-W)^m$, where m lies between 3/2 and 2, as is shown in fig. 2 for the experiments of Wooten. The corresponding functions I(E)

Fig. 2.



are then of the form $I(E) = \text{const.}\,\frac{(E-W)^{m-1}}{V^2}$. For the present case the variability of the mass of the electron with speed will not make any essential difference. Putting $\frac{1}{2}mV^2 = E$ we find that I(E) will increase rapidly to a maximal value at $E = \frac{W}{2-m}$, after which it decreases with increasing E. This is a characteristic feature exhibited by all ordinary ionization curves. However, the experimental material presently at hand covers so small a range of

* Webster and Clark, Proc. Nat. Acad. Sc. iii. p. 181 (1917).
† B. A. Wooten, Phys. Rev. xviii. p. 71 (1919).

voltage that even the existence of a maximum in I(E) cannot be inferred with certainty. It would therefore be very interesting to have the experiments extended to other elements and to a wider range of voltage covering also the gradual falling off of the ionization behind the maximum value.

In the main the function I(E) may, apart from a difference of linear scale, be expected to exhibit the same qualitative properties independent of the element and the X-ray lines which have been employed in the experiment. In detail, however, differences may occur, as for instance in the relative height and position of the maximal point. For values of E which are large compared with the ionization energy of the electron in the group in question I(E) may be expected to be given by the simple expression (3). The expression I(E) may in general be considered to be the product of the number n of electrons, contained in the group in question, and an individual ionization function which principally will depend upon the energy of the incident particle and the ionization potential of the electron in question, and only to a minor extent on the special kind of the interaction with the other electrons in the group. It might therefore be expected, that experiments of this kind will afford a direct manner of determining the number of electrons in the different electronic groups of the atom.

On the whole, experiments on characteristic X-ray intensity seem to offer a method of studying the process of ionization under various conditions. This method seems to possess several advantages, since each X-ray line is associated with the primary ionization of a definite electronic group, while in ordinary experiments on ionization in gases we have, apart from hydrogen and helium, to do with more complex phenomena. The experiments can further be extended to a much wider range of ionization potentials, and finally the part of the ionization curve accessible by this method is just the part which offers the greatest experimental difficulties in the case of experiments on ionization in gases.

§ 3. *Efficiency in producing characteristic X-rays.*

As we do not know the absolute value of the ionization performed by electrons in inner atomic groups, we cannot calculate the absolute value of the intensities of X-ray lines. An upper limit for the intensity should be expected to be

obtained by introducing I(E) from formula (3) in (15) and (16). Neglecting the relativity modifications we obtain

$$F = \frac{nz\pi e^4}{k}\left\{\frac{E_0 - W}{W} - \log\frac{E_0}{W}\right\}\sum_k a_{ik}h\nu_{ik}. \quad . \quad (19)$$

In the experiments referred to above the absolute value of the intensity is in no case recorded. An estimate of the ratio of the intensity of the X-ray lines to that of the continuous radiation in the case of molybdenium and palladium can be obtained from the measurements of Clayton Ulrey *. From Ulrey's diagrams it appears that this ratio for the characteristic K-lines is about 1/2 for Mo and 1/4 for Pd. The voltage applied was 35,000 volts.

The efficiency in producing general radiation can be calculated from the formula

$$\eta = 2{\cdot}7\,.\,10^{-10}AV, \quad V = \text{voltage}, A = \text{atomic weight}. \quad . \quad (20)$$

which is shown by Beatty † and Rutherford and Barnes ‡ to hold for a wide range of voltages and for a large number of elements. The constant factor corresponds to that found by Rutherford and Barnes §. Proceeding in this way we find for the efficiency in producing characteristic X-rays in Ulrey's experiments, $6{\cdot}5\,.\,10^{-4}$ and $3{\cdot}5\,.\,10^{-4}$ for molybdenum and palladium respectively.

From formula (19) we can calculate an upper limit for the efficiency F/Ez. The sum $\sum_k a_{ik}h\nu_{ik}$ will be of the order of W_i, while k can be interpolated from Whiddington's ‖ experiments on the decrease of velocity of cathode rays in Al and Au. The value adopted is $k = 3.10^{-35}$ gr.^{2} sec.$^{-4}$ for both Mo and Pa. With the values of the critical potentials found by Wooten and Ulrey the efficiency calculated from (21) is about 12.10^{-4} and 6.10^{-4} for Mo and Pa respectively, when assuming $n = 2$ for the K-group. These values are twice as large as those deduced from experiment, but the difference is probably not larger than was to be expected from the nature of the calculation, and the large experimental errors involved in the determination of the efficiency of producing general X-radiation as well as in the determination of the stopping power of the different elements for cathode rays.

* C. Ulrey, Phys. Rev. xi. p. 401 (1918).
† Beatty, Proc. Roy. Soc. lxxxix. p. 314 (1913).
‡ Rutherford and Barnes, Phil. Mag. xxx. p. 361 (1915).
§ Compare also Rutherford and Barnes, *l. c.* p. 367, footnote.
‖ Whiddington, Proc. Roy. Soc. lxxxvi. p. 360 (1912).

§ 4. *Enhanced lines in X-ray spectra.*

As mentioned above, Wentzel has pointed out that some X-ray lines which do not fit in the ordinary scheme of classification can be explained on the assumption that they are emitted by atoms which have lost more than one electron from an inner group. Wentzel supposes that this double ionization of the atom is due to successive impact. On account of the probably very short life-period of the excited state of an inner electronic group, this assumption can hardly be maintained. It will appear that the multiple ionization in question can to a measurable extent only be due to simultaneous action by a single impacting particle.

Consider the probability of successive collisions in the simple case when an ionized atom having lost a single electron from an inner group loses a second electron from the same group before reorganization of the atom has taken place. The probability that the ionized atom in question in the time dt will experience collisions with impacting particles resulting in a second ionization of the same group is

$$dF = n\pi {p_W}^2 \frac{s}{e} dt, \quad . \quad . \quad . \quad . \quad (21)$$

where n is the number of electrons in the group considered, s the current density of impacting particles, and p_W the radius of ionization. The probability that the atom in the time dt will suffer a spontaneous transition from the state considered under emission of radiation is $A\,dt$, where A is a constant. The ratio of the number of doubly ionized atoms to the number of singles is

$$\tau = \frac{\pi n {p_W}^2 s}{Ae}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (22)$$

This ratio will be seen to be vanishingly small for all cases when electrons are removed from inner groups of the atom where the probability of spontaneous transition is conditioned only by the state of the atom itself. Considerations based upon the principle of correspondence indicate that A ought to be of the order $10^8 . Z^4$, where Z is the effective nuclear charge of the group. On the other hand, ${p_W}^2$ will be of the order $10^{-16} Z^{-4}$. For a current density of 1 ampere s/e is of the order 10^{19}, and consequently it is seen that τ is very small in all cases involving the ionization of inner groups. Similar considerations apply also when the collisions do not result in the removal of an electron from an inner group,

but only in a transference of an outer electron from its normal orbit to an orbit characterized with a quantum number some units higher. This problem has been discussed by Compton*. As the life period of these states must increase rapidly with increasing quantum number, the part played by successive collisions will also increase rapidly, until when complete ionization of an outer group is obtained it may under favourable conditions play a preponderant part. While thus successive impact may be the essential cause of the multiple ionization responsible for the appearance of the optical spark spectra, we must assume that the multiple ionization giving rise to enhanced lines in X-ray spectra must be due to simultaneous action by a single impact.

The double ionization performed by α particles in helium has been shown to be intimately connected with the dimensions of the electronic orbits and the number of electrons in the helium atom. In the same manner the intensities of enhanced lines in the X-ray spectra are directly connected with the dimensions of each electronic group and the number of electrons contained in it. As no systematic measurements of this kind have been performed, we shall be content with showing in broad outlines that the experimental material at hand is consistent with the assumption that the multiple ionization is due to single impacts.

Assume for simplicity, as before, that the electrons are independently distributed on the surface of spherical shells. This special form of the distribution will not seriously influence the result, as an alternative assumption of uniform volume distribution within a sphere will lead to practically the same dependence upon r and p apart from a slowly varying factor. Consider the general case when an incident electron removes m_1 electrons from a group containing n_1 electrons . . . and m_s electrons from a group containing n_s electrons. Using the same notation as in § 2, section I., we obtain in analogy to the formulæ (4) and (5) for the probability of a particular arrangement of the electrons when s groups are concerned,

$$P=\prod_{i=1}^{i=s} \frac{n_i!\ \sigma_i^{m_i}}{m_i!\ (n_i-m_i)!}(1-\sigma_i)^{n_i-m_i}, \quad . \quad . \quad . \quad (23)$$

* Compton, Phys. Review, xv. (2) pp. 130 and 476 (1920). A significant exception occurs when the state in question is of the metastable kind, where A is equal to zero. In this case successive collisions with other atoms of the gas or with the impacting particles will necessarily occur. The importance of this point is mentioned by J. Franck, *Phys. Zs.* xxii. p. 443 (1921).

and for the number of collisions of the type in question experienced per unit path by the incident electron

$$I_m = 2\pi N \int P q dq, \quad . \quad . \quad . \quad . \quad . \quad . \quad (24)$$

where the integration is to be extended to all values of P different from zero. The expression for the intensity of a particular line with frequency ν_{ik} will be of the form

$$F_m = \frac{a_{ik} i h \nu_{ik}}{k N} \int_{W_c}^{E_0} I_m E dE. \quad . \quad . \quad . \quad . \quad (25)$$

As before, a_{ik} is the relative probability of the transition giving rise to the line in question, and W_c is the critical potential of the line. This potential will be different from the critical potential of the corresponding line, which is due to single ionization, and roughly equal to the sum of the potentials necessary to remove each electron separately from the neutral atom. According to (24) the intensity of enhanced X-ray lines will be expected to increase according to a law different from that which is valid for ordinary lines.

Consider for the sake of illustration the simple case of double ionization of the K-group. As the ratio $\frac{p}{r}$ in the case of elements with higher atomic numbers will be small, we can use the simple expression for I_2 given by (9) and neglect $\frac{p_2^2}{r^2}$ in the logarithmic term as well as the distinction between p_1 and p_2. Introducing further p^2 from (2) we thus obtain

$$F_2 = a_{ik} i \frac{h\nu_{ik}}{k} \frac{\pi}{8} \frac{e^8}{r^2 W^2} \int_{W_c}^{E_0} \log_n \left(\frac{r^2 E^2 W}{4e^4 (E - W)} \right) \times \frac{(E-W)^2}{E^3} dE. \quad . \quad . \quad . \quad (26)$$

The critical potential W_c is approximately equal to $2W$.

The relative intensity of an enhanced line to the corresponding line due to single ionization is seen, by using the expressions (19) and (26) for a constant ratio E_0/W, to be sensibly inversely proportional to $(2rW)^2$, which in this case will again be nearly equal to Z^2e^4. The enhanced X-ray lines in the K-series due to double ionization of the K-group may thus be expected to decrease rapidly with increasing atomic number of the element.

A similar conclusion will hold for the relative intensities of the other enhanced lines in the K- as well as in the other series, which, as may be inferred from (24), will decrease

with increasing Z. Further, as may be seen from (22), the relative intensities will decrease with increasing value of s.

These considerations are in accordance with experimental facts. First, the strongest enhanced lines observed correspond to double or triple ionization, and the lines which are due to double ionization are stronger than those due to triple ionization. Secondly, the intensity of enhanced lines decreases rapidly with increasing atomic number, so that in the K-series these lines have not been observed systematically for elements with atomic number higher than 30. It is further shown in a paper by Coster * which has just appeared, that the intensity of some lines in the L-series, which probably belong to the enhanced lines, increases in a different way with increasing voltage from the ordinary L-series lines and in a manner which qualitatively conforms to (26), although the measurements did not give an accurate measurement of the critical potential. This is of great interest in this connexion, and measurements especially on the critical voltage of the ordinary enhanced lines will afford a crucial test of this theory and a reliable means of disentangling the lines due to multiple ionization from those due to single ionization.

In the K-series in the elements from sodium to zinc one of the strongest enhanced lines observed is K_{α_1}, which corresponds to a transition of an L-electron into the doubly ionized K-level. Accurate measurements on the intensity of this line compared with that of the principal K_α lines for different voltages may, as seen from equation (25) or (26), offer direct test of the theoretical values for the dimensions of the K-group. In the same manner measurements on the intensity of the enhanced line L_{α_3} in the L-series may yield information about the dimensions of the L-group, and so on. It is, however, to be remembered that lines resulting from different processes of ionization may have the same frequency. Consider thus the line K_{α_2}, which appears when an electron from the singly ionized L-group passes into the singly ionized K-group. The state in which one electron is missing from the K-group and one from the L-group can, however, either be produced directly by a collision or indirectly from a state in which both K-electrons are removed.

In the same paper Wentzel, in support of his assumption that the enhanced X-ray lines were due to collisions by successive impact, refers to the complex structure of the absorption edges, observed by Fricke, Hertz, and Stenström,

* D. Coster, Phil. Mag. xliv. p. 546 (1922).

which he explains as due to absorption of atoms having already lost an electron by a previous process of absorption. This assumption, however, can scarcely provide the right explanation of the phenomenon. The relative intensity of the different "components" of an absorption edge must, namely, on this hypothesis be expected to be nearly equal to the ratio of the number of atoms with electrons absent from some inner group to the total number of neutral atoms in the material used in the experiment; but under the ordinary experimental condition this ratio will probably be far too small to yield any satisfactory explanation. The problem is examined in detail in the above-mentioned paper by Coster, where also a different explanation of the complex structure of the absorption edges is discussed.

Summary.

The theory of ionization by swiftly moving electrical particles is discussed in its relation (1) to the experiments of Millikan and Wilkins on the production of divalent ions in helium by α-rays, (2) to the increase of intensity of characteristic X-rays with increasing voltage, and (3) to the occurrence of certain lines in the X-ray series which are due to multiple ionization of the inner electronic groups (enhanced lines).

It is pointed out that the phenomenon of multiple ionization of an atom produced by the action of a single particle is directly connected with the dimensions of the different electronic groups. In this manner the dimensions of the electronic orbits in helium are calculated from Millikan's experiment, and found to be of the order predicted by Bohr's theory. It is also shown that experiments on the dependence of intensity of characteristic X-rays on voltage afford a direct means of investigating the number of electrons in the inner electronic groups of the heavier atoms and the dimensions of these groups. In this connexion the appearance of X-ray lines originating from doubly ionized atoms is discussed, and it is concluded that these lines are excited by a single impact of an electron on the atom and not by successive impacts.

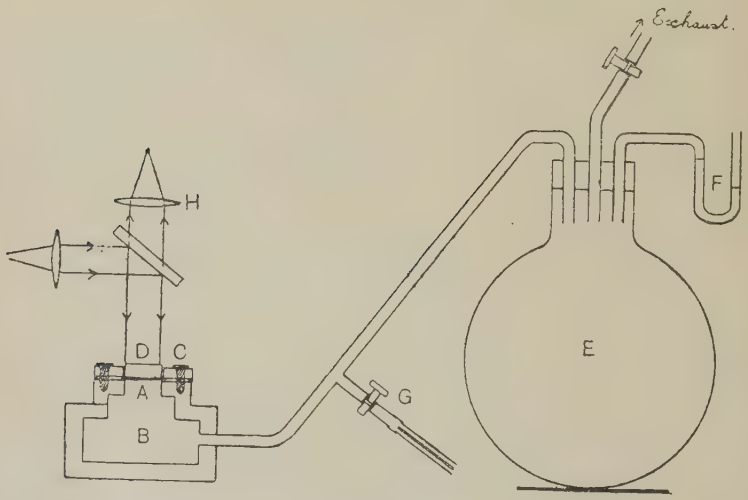
In conclusion I wish to express my best thanks to Prof. N. Bohr for his kind interest and help throughout the work.

Institute of Theoretical Physics,
University of Copenhagen.

IV. *Application of an Optical Manometer to the Measurement of the Viscosity of Gases.* *By* J. E. P. Wagstaff, *M.A., Lecturer in Physics at the University of Leeds* *.

A DELICATE form of optical manometer, recently described in the Proceedings of the Camb. Phil. Soc.†, has been used as a pressure-gauge for measurements of the viscosity of air. This method of measuring small changes of pressure has been found to possess considerable advantages over the usual method adopted in this type of experiment of taking two simultaneous readings of a moving mercury column. A thin slip of glass A (diameter 2 cm.), whose modulus of elasticity is known, is made to rest on the top

Fig. 1.



of a pressure-chamber B (*vide* fig. 1), in the upper surface of which is a circular hole of diameter 1·93 cm. In order to clamp the edges of the slip, a brass plate C, having a hole of diameter 1·97 cm. concentric with that in the upper surface of the chamber, is screwed down on the top surface of the glass slip. A small plate of glass D (diameter 1·95 cm. approx. and thickness 1 mm.) is made to rest on the top of the slip, care being taken that the air in the space between

* Communicated by the Author.

† Wagstaff, Proc. Camb. Phil. Soc. vol. xx. Part 1.

the slip and the plate remains at atmospheric pressure. The chamber is connected to a reservoir E, which can be partially exhausted by means of a filter pump. Air can be allowed to enter the apparatus very slowly through a fine capillary tube of known dimensions by opening the tap G. The film between A and D is illuminated by a parallel beam of monochromatic light, and is viewed by a microscope at H. The time taken for a given number of circular fringes to close in towards the centre is observed, from which the change in pressure in the chamber E in a given time can be calculated, and the viscosity determined.

Theory. In the case of a gas of viscosity η flowing through a tube of length l cm. and radius a cm., the rate of flow is determined by the equation

$$p_1 v_1 = \frac{\pi a^4 (p_2^2 - p_1^2)}{16\eta l},$$

where p_1 and p_2 are the pressures of the gas at the two ends of the tube, and v_1 is the volume of the gas flowing through the tube per sec. at pressure p_1.

If S = combined volume of the vessel E and of the pressure-gauge B, p_2 the pressure of the external atmosphere, and T the time taken for the pressure inside the apparatus to rise from p_1 to p_3, it has been shown that *

$$\frac{1}{2p_2}\left[\log_e \frac{p_2+p}{p_2-p}\right]_{p_1}^{p_3} = \frac{\pi a^4 \mathrm{T}}{16\eta l \mathrm{S}}.$$

Write $p_3 = p_1 + p_1'$,

then $$\log_e\left(1+\frac{p_1'}{p_1+p_2}\right) - \log_e\left(1-\frac{p_1'}{p_2-p_1}\right) = \frac{\pi a^4 \mathrm{T} p_2}{8\eta l s},$$

where p_1', corresponding to the disappearance of a given number of fringes for light of wave-length λ, can readily be calculated from the formula

$$p_1' = \frac{256 \mathrm{E} t^3 \mathrm{N} \lambda}{90 \mathrm{R}^4},$$

for a plate of radius R, thickness t, and Young's modulus E, which is clamped at the edge. Several sets of observations were made for various values of p_1, the initial pressure inside the apparatus, which can be observed quite accurately with a cathetometer ; the results are tabulated below.

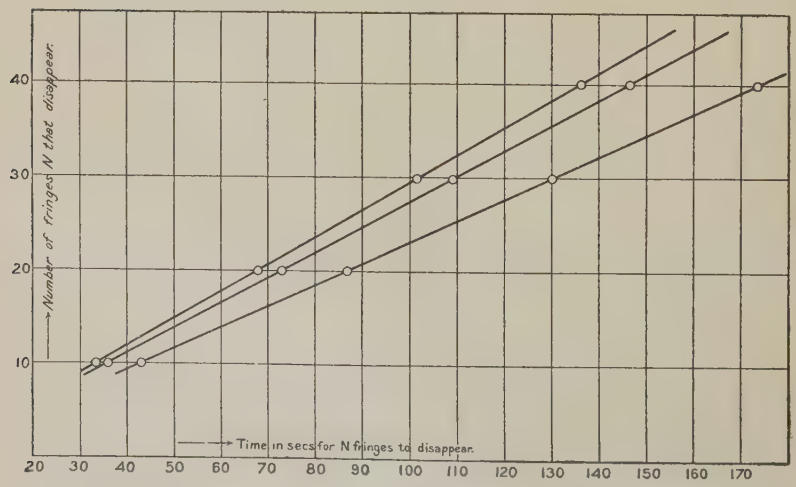
* Searle, Camb. Phil. Soc. vol. xvii. Part 2, p. 183.

Series I.

Length of capillary tube = 61·6 cm.
Barometric pressure = p_2 = 75·89 cm.
Temperature = 18°·5 C.

$p_2 - p_1$.	N = number of fringes that disappear.	T = time taken for N fringes to disappear.
24 cm.	10	33·6 sec.
	20	68 ,,
	30	101·5 ,,
	40	136·5 ,,
22·1 cm.	10	36 ,,
	20	73 ,,
	30	109·3 ,,
	40	147 ,,
18·1 cm.	10	43 ,,
	20	87 ,,
	30	130·5 ,,
	40	173·9 ,,

Fig. 2.


Number of fringes N that disappear.
40
30
20
10
Time in secs for N fringes to disappear.
20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170

When N is plotted against T as in fig. 2, a straight line relationship is indicated.

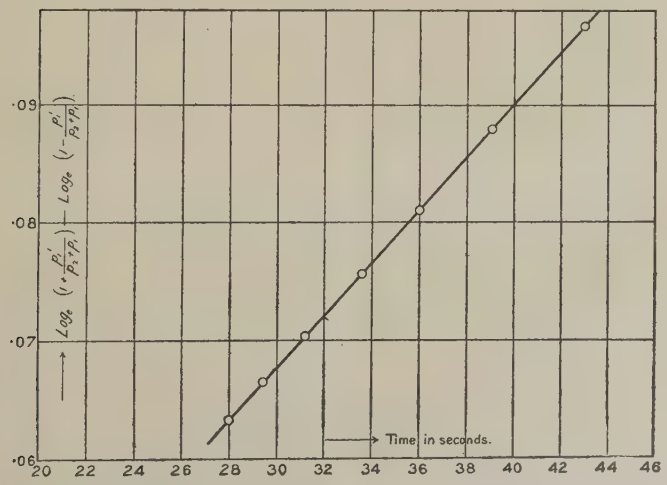
In a second series of experiments, the time taken for a given number of fringes to disappear for different initial pressures is observed ; the results are given below.

SERIES II.

Volume of reservoir E and chamber B = 4,800 c.c.
$\theta = 18°{\cdot}5$ C. N = 10.

$p_2 - p_1$ cm. of mercury.	Time taken for 10 fringes to disappear.	$p_1' = {\cdot}149$ N.	$\log_e\left(1+\frac{p_1'}{p_2+p_1}\right) - \log_e\left(1-\frac{p_1'}{p_2-p_1}\right)$.
30·0	28·0 sec.		·06310
28·0	29·5 „		·06631
26·1	31·3 „	1·49	·07023
24·0	33·6 „		·07565
22·1	36·0 „		·08104
20·1	39·0 „		·08818
18·1	43·0 „		·09670

Fig. 3.



Columns 2 and 4 of the above table have been plotted in fig. 3, and a linear relationship between the quantities is indicated. The best fitting straight line through the observations

gives a mean value for

$$\frac{\log_e\left(1+\frac{p_1'}{p_2+p_1}\right)-\log_e\left(1-\frac{p_1'}{p_2-p_1}\right)}{\mathrm{T}}=\frac{1}{440{\cdot}6}.$$

The radius of the tube is found in the usual way by inserting a thread of mercury and finding its length and mass. The mean value of the viscosity of air, under the conditions of the experiment, is now easily deduced and is found to be $1{\cdot}811\times10^{-4}$ dyne per cm.^{2} per unit velocity gradient. This is in very close agreement with determinations by other methods.

Finally I wish to thank Mr. A. Hare, B.Sc., for his assistance while carrying out the investigation.

V. *The Shape of the Capillary Curve formed between a Flat Glass Plate and a Uniform Circular Cylinder, and the Accurate Determination of the Surface Tension of Liquids.* *By* J. E. P. Wagstaff, *M.A., Lecturer in Physics at the University of Leeds* *.

Introduction.

IN the determination of surface tension by means of a capillary tube, one of the difficulties is the accurate measurement of the diameter of the tube at the point to which the liquid rises. Such measurements are always open to considerable inaccuracy. In the following paper I have attempted to eliminate this source of error by forming the meniscus between a uniform circular cylinder containing liquid, and a plate dipping into the liquid and cutting the cylinder along two generators. The problem then resolves itself into the determination of the diameter of a wide tube (say 5 cm.) and of the width of a plate (say 2 cm.). Moreover, the method permits of a number of observations being made on the same liquid during one experiment, and from these the value of the surface tension can be accurately determined.

Description of Experiment.

A plate and a thin-walled circular cylinder, of uniform section, both of glass, are first carefully cleaned by washing with potassium dichromate and concentrated sulphuric acid,

* Communicated by the Author.

and rinsing with distilled water. The cylinder, containing the plate, is then placed on a levelling-table. The liquid under observation is then introduced into the cylinder and the plate is pressed against the wall of the latter so that its edges are vertical, by means of two small glass rods with rubber tips inserted horizontally between the plate and the opposite wall of the cylinder. This is insured by viewing the plate through a microscope with a vertical travel, making a scale-division in the eyepiece coincide with the edge of the plate, and adjusting the screws attached to the table until the scale-division coincides with the edge at all heights. The measurements are made by means of a travelling microscope having a vertical movement which rests on the bed of a measuring machine capable of horizontal motion.

Theory.

Let the curved surface ABEDCF be a portion of the cylindrical vessel of radius R, and ABCD the glass plate of width $2a$ which intersects the cylinder along the generators AD and BC, and suppose that the bottom of the meniscus of the liquid confined between the plate and the cylinder assumes the shape QSPR as indicated in fig. 1. Take a system of rectangular axes at a point O, the intersection of the line GH joining the mid points of AB and CD with the level of the surface of the liquid in the outer vessel, such that OY lies along HG, and OX is in the plane of the plate. Supposing the principal radii of curvature at $P(x_1y_1)$ are R_1 and R_2, R_2 being parallel to the plane of the plate, the difference of pressure p between the two sides of the meniscus is given by

$$p=y_1\rho g=T\left(\frac{1}{R_1}+\frac{1}{R_2}\right),$$

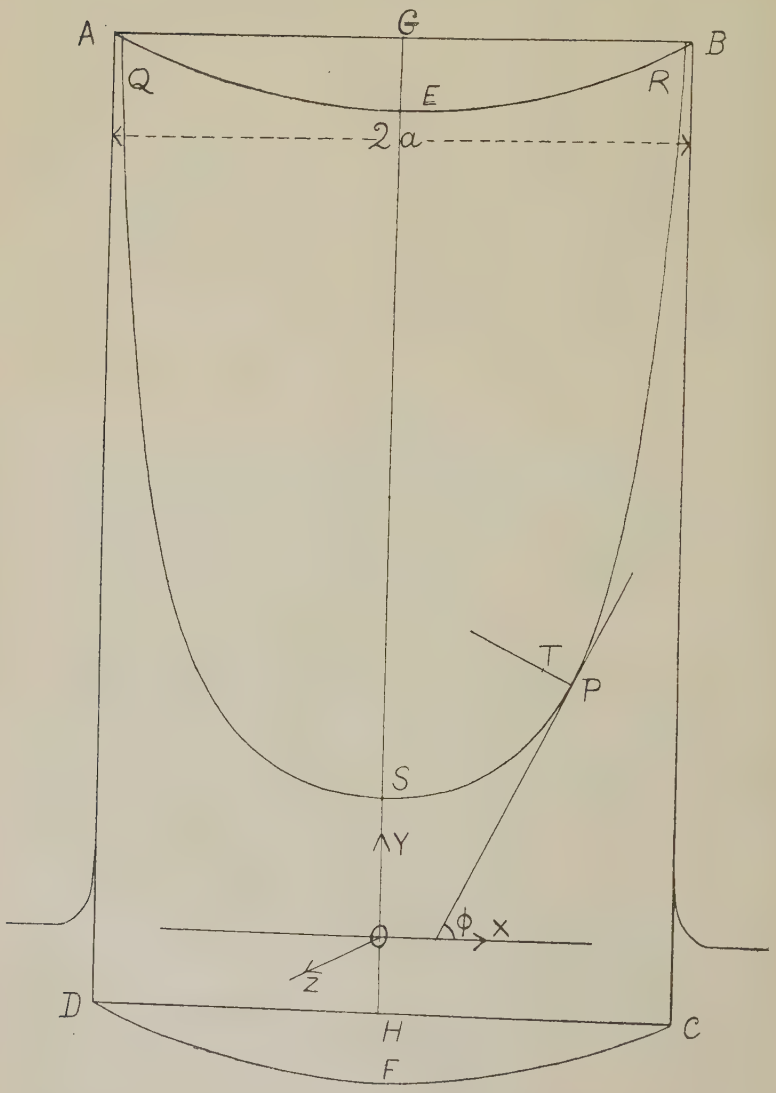
for a liquid of density ρ and surface tension T. Neglecting $\frac{1}{R_2}$ in comparison with $\frac{1}{R_1}$,

$$y_1\rho g=\frac{T}{R_1}.$$

The equation of the normal to the meniscus at P is

$$y-y_1=-\frac{x-x_1}{\tan\phi},$$

Fig. 1.



and the distance from P of any point T(x, y) on this line is

$$TP^2=(x_1-x)^2+(y_1-y)^2=\frac{(x_1-x)^2}{\sin^2\phi}. \quad . \quad . \quad . \quad \text{(I.)}$$

Let z_0 be the distance between the two glass surfaces at S, and z the corresponding distance at T, then

$$(z_0-z)(2R-z_0+z)=x^2, \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(II.)}$$

and

$$z_0(2R-z_0)=a^2. \quad . \quad . \quad . \quad . \quad . \quad \text{(III.)}$$

Moreover, if TP is the depth of the section of the surface of the meniscus at P, by a plane through TP drawn perpendicular to the plane of the figure,

$$TP=R_1(1-\sin\psi) \quad \text{and} \quad z=2R_1\cos\psi,$$

where ψ is the angle of contact between glass and the liquid. In the case of water and glass $\psi=0$.

Thus $(z_0-2R_1)(2R-z_0+2R_1)=x^2$ for $\psi=0$,

and

$$\frac{T}{R_1}=y_1\rho g.$$

Substituting for R_1, and writing $\frac{2T}{\rho g}=K$,

$$\left(z_0-\frac{K}{y_1}\right)\left(2R-z_0+\frac{K}{y_1}\right)=\left(x_1-\frac{K\sin\phi}{2y_1}\right)^2.$$

Inserting

$$\frac{K}{z_0}=y_0,$$

$$K\left(\frac{1}{y_0}-\frac{1}{y_1}\right)\left(2R-K\left[\frac{1}{y_0}-\frac{1}{y_1}\right]\right)=\left(x_1-\frac{K\sin\phi}{2y_1}\right)^2,$$

which is the equation of the projection of the bottom of the meniscus between the two surfaces on the plane $z=0$. This equation suggests a method of determining T. The quantities $K\left(\frac{1}{y_0}-\frac{1}{y_1}\right)$ and $\frac{K\sin\phi}{2y_1}$ are small in comparison with 2R and x_1 respectively, so that an approximate value of K will be sufficient to estimate these terms. This approximate value of K is given by

$$K=z_0y_0,$$

where $z_0(2R-z_0)=a^2$.

The angle ϕ is measured directly from the graph showing x_1 for different values of y_1 (*vide* fig. 2). In practice $2x_1$ is observed at different heights y_1. If now, in the case of any

liquid,

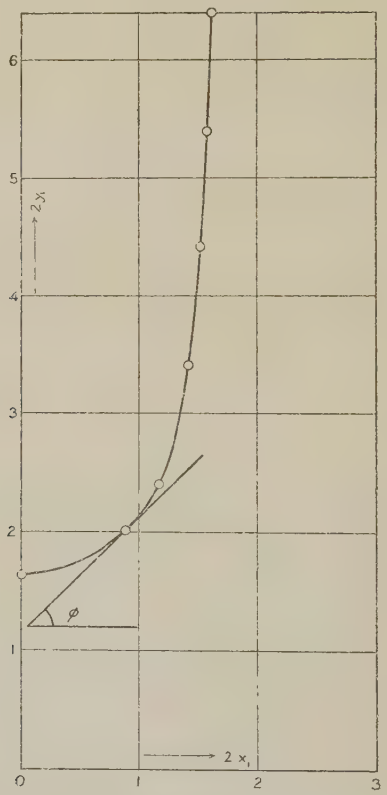
$$\left(\frac{1}{y_0}-\frac{1}{y_1}\right)\left(2R-K\left[\frac{1}{y_0}-\frac{1}{y_1}\right]\right)$$

be plotted against

$$\left(x_1-\frac{K\sin\phi}{2y_1}\right)^2,$$

the points obtained are found to lie on a straight line as indicated by the theory, and the slope of this line gives the surface tension.

Fig. 2.



Several sets of observations have been taken for different liquids. Tabulated below are the results obtained with distilled water at 18° C.

Table I.

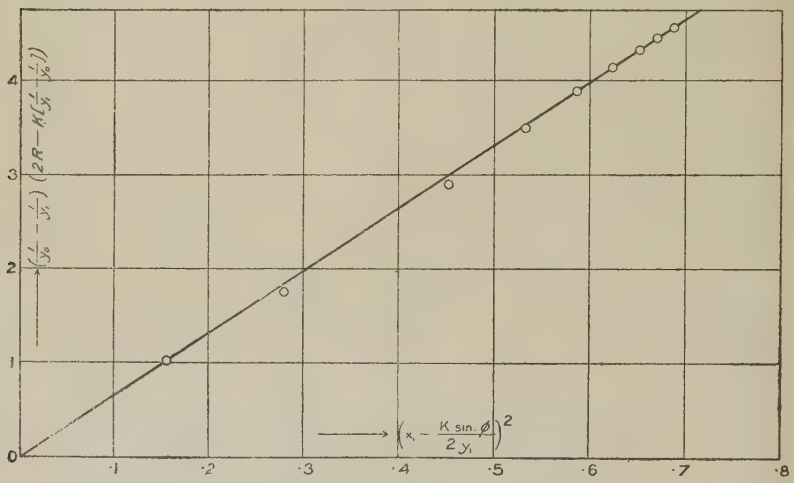
$2a = 17{\cdot}6765$ mm. $\quad 2R = 4{\cdot}726$ cm. $\quad y_0 = {\cdot}825$ cm.

Approximate value of $K = 1/7$. $\quad$ Temperature of water = 18° C.

$2x_1$ (cm.).	x_1 (cm.).	y_1 (cm.).	$\frac{y_1-y_0}{y_1y_0}$.	$K\frac{y_1-y_0}{y_1y_0}$.	$2R-K\frac{y_1-y_0}{y_1y_0}$.	$\frac{y_1-y_0}{y_1y_0}\left\{2R-K\left(\frac{y_1-y_0}{y_1y_0}\right)\right\}$.	ϕ.	$\frac{K\sin\phi}{2y_1}$.	$x_1-\frac{K\sin\phi}{2y_1}$.	$\left(x_1-\frac{K\sin\phi}{2y_1}\right)^2$.
0	0	$y_0={\cdot}825$	0	0	4·726	0	0	0	0	0
·87898	·43949	1·005	·2171	·03101	4·695	1·020	40°	·04568	·39381	·1551
1·16583	·58292	1·205	·3823	·05461	4·6714	1·785	67°	·05458	·52834	·2790
1·4259	·71295	1·705	·6258	·08940	4·6366	2·902	80°	·04128	·67167	·4512
1·5265	·76325	2·205	·7586	·10837	4·6176	3·502	85°	·03227	·73098	·5343
1·58498	·79249	2·705	·8424	·12034	4·6057	3·881	87°	·02637	·76612	·5869
1·62388	·81194	3·205	·9001	·12859	4·5974	4·137	87·5°	·02226	·78968	·6237
1·65198	·82599	3·705	·9421	·13459	4·5914	4·325	88°	·01927	·80672	·6507
1·66808	·83404	4·205	·9743	·13919	4·5868	4·470	89·5°	·01699	·81705	·6676
1·68613	·84307	4·705	·9993	·14276	4·5832	4·581	89 5°	·01518	·82789	·6855
1·69725	·84862	5·205	1·020	·14571	4·5803	4·672	89·5°	·01373	·83489	·6972

If θ is the angle of inclination of the straight line drawn through the observations (fig. 3),

Fig. 3.



$$\cot\theta = \mathrm{K} = \frac{2\mathrm{T}}{\rho g}.$$

Hence the surface tension of water at 18° C. = 74 dynes per cm.

VI. *A Critical Review of Trouton's Law, and its Applicability at the Triple Point. By* Sasi Bhusan Mali, *M.Sc., Research Assistant, University College of Science, Calcutta*.*

ACCORDING to the well-known law propounded by Trouton, the ratio of the molecular latent heat of vaporization of a substance to the temperature of vaporization is a constant for all substances, and the value of the constant is approximately equal to 21. It was shown later that the Trouton law is applicable not only in the case of vaporization but also in the case of fusion of substances, the value of the constant being of course different in the latter case. Hence, the generalized Trouton law may be expressed by saying that the latent heat of change of state of a substance

* Communicated by Prof. Megh Nad Saha.

is proportional to the temperature of the change of state, and the ratio of the two is approximately the same for all substances provided the same kind of change of state is involved.

By considering the case of a large number of simple substances, Nernst* has shown that the original Trouton law for vaporization does not hold generally, and he therefore connects the latent heat of vaporization with the temperature of vaporization by an empirical formula of the type

$$\frac{M\lambda}{\theta} = A \log \theta - B\theta,$$

where A and B are constants.

The object of the present paper is to examine critically the validity, and the theoretical foundation, if any, of the Trouton law. It is obvious that the quotient $\frac{M\lambda}{\theta}$ denotes Molecular Change of Entropy during a discontinuous physical change. Thus, if we confine our attention to the phenomena of vaporization, we have

$$\frac{M\lambda}{\theta} = S_1 - S_2,$$

where S_1 = gram-molecular entropy of the vapour phase, and S_2 = gram-molecular entropy of the liquid phase.

The value of the quotient is generally calculated at the boiling-point, *i. e.*, at the temperature at which the substance boils under a pressure of one atmosphere. Now, this particular pressure is a totally artificial point. It has no physical significance such as the triple point, the critical point, or the inversion point has. It varies enormously with the external pressure, while the latent heat varies neither in the same direction nor to the same extent. Thus, let us take the case of water. We calculate the Trouton quotient for water boiling under different pressures.

P.	B.P.	Latent Heat.	Trouton Quotient.
4·6 mm.	0° C.	606·5	40
760 mm.	100° C.	535·9	25·9
20950 mm.	230° C.	446	16

* W. Nernst, *Wärme Satze*, p. 117.

At the critical point the Trouton quotient becomes zero, because the latent heat vanishes there. Thus, the Trouton quotient varies from 0 to 40 as we pass the range downwards from the critical point to the triple point.

An attempt has been made to show that the Trouton quotient may have the same value for all substances, when we calculate it at the corresponding temperatures of these substances. Guldberg * and Guye † tried to show that the boiling-point under atmospheric pressure is two-thirds of the critical point for many substances, which, if true, would afford a theoretical basis for the Trouton law. But unfortunately the relation is very rough as the following table shows.

Substance‡.	Critical point in absolute temp.	Boiling-point in absol. temp.	Quotient.
Ethyl ether	466·8	307·6	0·66
Ethyl alcohol	516·1	351·4	0·68
Ammonia	405·53	240	0·59
Argon	152	86·9	0·57
Benzol	561·5	353·2	0·63
Bromine	575·2	331·6	0·58
Hydrobromic acid	364·3	204·5	0·56
Chlorine	419	306·7	0·73
Chloroform	531·8	334·2	0·63
Hydrochloric acid	324	190	0·59
Carbon dioxide	304	193·8	0 64
Methyl alcohol	513	337·7	0·66
Oxygen	155	90·1	0·58
Sulphuretted hydrogen	374·4	213	0·57
Sulphur dioxide	430	263	0·61
Nitrogen	126·5	77·33	0·61
Water	647	373	0·58
Hydrogen	32	20·22	0·63

It appeared to the present writer that the Trouton quotient might be found to be constant at the triple point. But calculations which are given below do not support this view. The triple point is only slightly different from the melting-point, therefore the data at the latter point have been used throughout the calculations. We can calculate the latent heat at the melting-point from that at the boiling-point with the aid of these data, making use of Planck's equation

$$\frac{dL}{dT} = (C_{P_1} - C_{P_2}) + \frac{L}{T} - T\frac{dP}{dT}\frac{\delta(V_1 - V_2)}{\delta T}.$$

* Guldberg, *Zeits. Physik. Chem.* v. p. 374 (1890).
† Guye, *Bull. Soc. Chim.* iv. p. 262 (1890).
‡ Landolt and Börnstein, *Tabellen*, pp. 439, 190, and 373.

The last term on the right-hand side is approximately equal to $\frac{L}{T}$ provided we neglect V_2 in comparison with V_1 and apply the perfect gas laws. Then

$$\frac{dL}{dT} = C_{P_1} - C_{P_2},$$

and

$$L_M = L_B - (C_{P_1} - C_{P_2})(T_B - T_M) \text{ approximately,}$$

where

L_M = gram-molecular latent heat at melting-point,
L_B = gram-molecular latent heat at boiling-point,
C_{P_1} = gram-molecular specific heat of the vapour phase,
C_{P_2} = gram-molecular specific heat of the liquid phase,
T_B = temperature of boiling under one atmosphere,
T_M = temperature of melting.

The results of calculation for a few simple substances are given in the following table :—

Substance*.	L_B.	C_{P_1}.	C_{P_2}.	T_B.	T_M.	L_M.	$\frac{L_M}{T_M}$.
Nitrogen	1363	6·4	13·3	77° K.	62°·5 K.	1434	23
Hydrogen ...	229	3	4·37	20°·4 K.	14° K.	237·8	17
Argon	1501	3	10·5	87° K.	85° K.	1516	18
Oxygen	1640	7	12·74	90°·1 K.	54°·23 K.	1846	34
Water	...	...	...	...	273° K.	10917	40
Bromine	6960	8·45	16·16	331°·6 K.	265°·68 K.	7462	28

It thus seems to be fairly well-established that there is no particular point of temperature at which the Trouton law holds. Neither does the law seem to have any theoretical significance. The fact that the law approximately holds good for a few selected substances when the vaporization considered is under one atmospheric pressure is of the nature of an accident, rather than of a law.

* Eucken, *Ber. d. Deut. Phys. Ges.* xviii. p. 4 (1916); W. Nernst, *Wärme Satze*, p. 117.
Landolt and Börnstein, *Tabellen*, pp. 185, 750, and 190.

VII. *Cathode Disintegration.* *By* The Research Staff of the General Electric Company, Ltd., London *.

Summary.

THE disintegration of a tungsten cathode struck by positive ions of various gases at various speeds has been investigated by using as the disintegrated cathode the "grid" of a triode.

(1) The rate of disintegration is independent of the temperature of the cathode.

(2) It increases regularly with the energy of the positive ions and is appreciable when the energy does not exceed greatly that required to ionize the gas. There is no irregularity in the neighbourhood of the normal cathode fall.

(3) It varies with the nature of the gas. The gases investigated could be arranged in the following order of increasing rate of disintegration :—Hydrogen, helium, nitrogen, neon, mercury vapour, argon.

No theory of disintegration is offered at present, for the results are professedly preliminary. But it is suggested that the facts as they stand throw some light upon unsettled problems of the unsupported electric discharge.

Incidentally some observations of the disappearance of the gas in the discharge are recorded; and it is pointed out how well the loss of mass of a wire may be measured by means of its decrease in conductivity.

Conclusions 1, 2, 3 confirm those of a more limited investigation described by Kingdon and Langmuir in a paper published since this work was completed (Phys. Rev. vol. xx. p. 107 (1922)).

Introduction.

1. Our knowledge of cathodic disintegration † is based chiefly upon experiments made on the unsupported discharge between cold electrodes. It has been established that with the same cathode in the same gas, the rate of disintegration is proportional to the current and increases with the cathode fall; so long as the cathode fall is "abnormal" the pressure of the gas is important only in so far as it affects

* Communicated by the Director.

† The work published before 1912 is summarized by V. Kohlschütter (*Jahrb. d. Rad. u. Elek.* vol. ix. p. 355 (1912)). References will therefore be only given to work subsequent to that date.

these two factors. In general the loss of mass with the same current and the same cathode fall increases with the molecular weight of the gas and the atomic weight of the metal. The relation between the disintegration and the electrical conditions obviously suggests that the action is due to the impact of positive ions on the cathode and increases with their speed; but since there is as yet no complete theory of the cathode fall, it is impossible to translate in detail information about the effect of change of cathode fall into information about the effect of change of speed. Further, there are certain anomalies shown by the inactive gases; argon and, to a lesser degree, neon are known to be peculiarly active in promoting disintegration when pure, while the disintegration is very greatly reduced if they are mixed with very small traces, less than 1 per cent., of the common gases *; even mercury is effective in reducing the disintegration, although its molecular weight is high. Now the inactive gases have peculiarly low normal cathode fall; it is uncertain whether their peculiarity in causing disintegration is due to a peculiar relation between speed of the ions and the cathode fall, or whether it is really a peculiarity in the disintegration due to ions of a given speed.

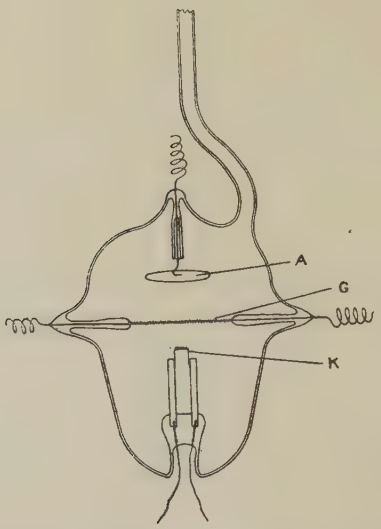
These difficulties in interpretation are made much less serious if a form of discharge is used for which a more complete theory is available. Such a discharge may be obtained in a triode with thermionic cathode. The current to the grid, when the grid is negative to the cathode, is then a measure of the number of positive ions arriving at it, and the speed of the ions arriving is related comparatively simply to the potentials of the electrodes. No experiments on the disintegration of the grid in such a vessel appear to have been undertaken. The measurements described in this paper show that it is a powerful method of investigating disintegration. The results presented here are merely preliminary, but they are sufficient to indicate the scope of the method and to show that the study of disintegration may throw some light on the still unsolved problems of the unsupported discharge. It is because they open up several lines of research in which a knowledge of the laws of disintegration could be used to investigate other problems—lines which we may not be able to follow up ourselves immediately—that we give an account of our experiments in this stage.

* J. Pintsch, Brit. Pat. 155783; D.R.P. Dec. 12, 1917.

Experimental Methods.

2. The discharge vessel employed in these experiments is shown in fig. 1. The cathode K is a tungsten spiral heated by a 6-volt battery, the anode A a nickel plate. (It would doubtless have been better to use a metallic coating on the inside of the bulb as anode, but as many observations had been made before this was realized, the apparatus was not changed.) The grid G is a close spiral of fine tungsten wire, such as is used in gas-filled lamps. The diameter of the wire was ·05 mm., of the spiral ·24 mm. The vessel was connected through a liquid air trap to a high vacuum pump,

Fig. 1



a McLeod gauge and apparatus for filling and purifying the gases used. A and G were connected to the outside terminals of a potentiometer, between which a potential V_0 was maintained by a battery or generator ; K was connected to the movable contact of the potentiometer so that its potential relative to A (V_a) and relative to G (V_g) could be varied.

A microammeter, shunted if necessary, could be inserted alternatively either in the grid or anode circuits, so as to measure the current i_g flowing to the grid, or the current i_a flowing from the anode ; i_g was from $0{\cdot}2$ to $2{\cdot}0 \times 10^{-6}$ amp. ; i_a from 15 to 200 times i_g.

3. To determine the loss of mass of the wire its resistance

was measured. This method has often been used before for determining the thinning of a wire *, but it does not seem to have been pointed out how satisfactory it is ; it is fairly accurate even if the diameter is not reduced by the same amount at all points along its length. If r is the radius of the wire at distance x from its middle point, ρ the density, σ the specific resistance, l the length, m the mass, R the resistance, and C the conductance, we have

$$m=\int_{-\frac{l}{2}}^{+\frac{l}{2}}\rho\pi r^2 dx\ ;\ \mathrm{R}=\int_{-\frac{l}{2}}^{+\frac{l}{2}}\frac{\sigma}{\pi r^2}dx. \quad . \quad . \quad . \quad (1)$$

If r is constant,

$$m\mathrm{R}=m/\mathrm{C}=\rho\sigma l^2. \quad . \quad . \quad . \quad . \quad . \quad (2)$$

If r varies, but varies symmetrically on both sides of the middle (as is the case here and in many other examples), we may write

$$r=r_0(1+ax^2+bx^4+\ldots), \quad . \quad . \quad . \quad . \quad (3)$$

and, as a first approximation, neglect powers of x higher than the second.

Then

$$\frac{m}{\mathrm{C}}=\rho\sigma l^2\left(1-\frac{a^2l^4}{576}\right); \quad . \quad . \quad . \quad . \quad . \quad (4)$$

so that if the correcting term in (3) for the extreme ends $\left(x=\frac{l}{2}\right)$ is as great as $\frac{1}{4}$, that in (4) is less than $\frac{1}{500}$. Accordingly when, as in these experiments, we are dealing with changes in resistance the greatest of which never reaches 25 per cent. of the original resistance of the wire, we may assume that $\frac{m}{\mathrm{C}}$ does not vary by more than 1 per cent., and write with that degree of accuracy

$$dm=m_0/\mathrm{C}_0\,.\,d\mathrm{C}, \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where m_0 and C_0 are the original mass and conductivity.

In order to be as free as possible from temperature variations, the grid was balanced in a bridge against a similar spiral in a similar vessel through which no discharge was ever passed. This spiral was taken throughout as the unit of resistance and conductivity ; its resistance at 18° C. was about 10 ohms, but it does not need to be known accurately. For if we measure in terms of this standard

* *E. g.* I. Langmuir, *Phys. Zeit.* xiv. p. 1273 (1913).

the conductivity C′ of various lengths of the wire of which the spirals are made and the corresponding mass m', we shall find $m'/C' = \text{constant} = \alpha$ (say). Then in (5)

$$m_0/C_0 = \frac{\alpha}{C_0^2}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

and

$$dm = \frac{\alpha}{C_0^2} \cdot dC, \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

whereby dm is determined in grams by means of measurements of conductivity in terms of the arbitrary standard. For the wire and the standard used in these experiments, α was ·0216 when m is expressed in grams. C_0 was nearly 3. (The standard spiral was about 3 times as long as the grids, being a survival from an earlier type of discharge vessel, which had been abandoned before these experiments were undertaken.) A change in conductivity amounting to 1/1000 part of the standard or 1/3000 part of the grid in its initial condition could be detected ; this corresponds to a loss of weight of about ·0024 mgm. The method is therefore sensitive as well as accurate.

Owing to the residual changes of conductivity with temperature which were not completely eliminated, the possible error in an estimation of the loss of mass increased with the duration of the experiment and with the consequent chance of change of temperature. An apparent decrease in conductivity of 1 in 1000 occurring during an experiment lasting an hour might be due to change of room temperature ; the shortest possible experiment lasted 20 minutes, which was the time that the vessel took to cool after being heated by the discharge. Accordingly there was a limit to the rate of disintegration which could be measured with accuracy, and the estimates of very slow rates, such as those characteristic of hydrogen, were unreliable. The values given for hydrogen or helium may possibly be in error by as much as 50 per cent. ; but since this source of error is absolute and not relative, the probable error of this nature with the remaining gases is not more than 2 or 3 per cent. It will be seen in the curves given below that the points vary from the mean line by a far greater amount than this. This variation is not mere experimental error in making the measurements ; the larger discrepancies are due to some real variation in the rate of sputtering, the cause of which is not yet completely understood.

The limit of accuracy of the conductivity measurements is set by the heating of the grid and standard by the measuring

current; when the discharge vessel is evacuated, a heating current of 1 ma. changes the resistance by 2 or 3 parts in a thousand. While the grid and standard are of the same diameter this effect is largely compensated, but after disintegration has proceeded for some time the compensation becomes imperfect. Another source of error lies in the leads which, in the arrangement adopted, had a resistance of about 1 per cent. of the standard. Since these various sources of error would have made the absolute values of the loss of mass uncertain to 5 per cent., and since none of the conclusions drawn depend upon absolute values, the work of reduction has been reduced by assuming C_0 the same for all grids. The absolute values of the loss may thus be in error by as much as 20 per cent., but the relative values in different experiments are correct to within 1 or 2 per cent.

4. As might be expected, all observations showed that, when the nature and pressure of the gas, the condition of the grid, and the potentials between the electrodes were unchanged, the loss of mass was proportional to the total quantity of electricity carried to the grid, *i. e.* $\int_T i_g dt$, where T is the time of the discharge. It is therefore convenient to express the disintegration as loss of mass per coulomb of grid current. An even more convenient method of expression is by number of atoms of tungsten lost per positive ion striking the grid. The disintegration S expressed in this manner is given by

$$S = \frac{e}{m_W} \cdot \frac{dm}{\int i_g dt}, \quad . \quad . \quad . \quad . \quad . \quad (8)$$

where m_W is the mass of the tungsten atom (at. wt. 184) and e the charge of an ion. Therefore

$$S = 523 \frac{dm}{\int i_g dt}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

But one source of uncertainty should be noticed. It is here assumed that all the current i_g is carried by positive ions; the current carried by the electrons liberated by the impact of the positive ions is neglected. Experiments are not available to estimate the fraction of the currents carried by these electrons, but there is good evidence that it is a small fraction. Cheney * found that the ratio of the electron to the positive ion current for a platinum cathode struck by Li, K, or Rb atoms with energies corresponding to a few

* W. L. Cheney, Phys. Rev. x. p. 335 (1917).

hundred volts never exceeded 0·1. The fraction apparently increased as the atomic weight of the metal struck decreased, but, since for Al it was only about twice that for Pt, it would not be much greater for W than for Pt. If the fraction were independent of the nature of the positive ions and of their energy, the only effect of the secondary electron emission would be to make the apparent value of S smaller than the real value in a constant ratio; but as this condition is not fulfilled, the relative values of S estimated from (9) may be somewhat in error. However, it seems legitimate to conclude that this source of error will not invalidate such conclusions, mainly qualitative, as are to be drawn here.

5. The adoption of the form of discharge vessel shown in fig. 1 does not wholly remove the uncertainty concerning the speed with which the positive ions strike the metal which they disintegrate; for the ions are produced throughout the volume of the vessel and the speed with which they reach the grid will depend upon the potential relative to the grid of the place where they are formed. But if the pressure of the gas is so low that the mean free path of the ions is greater than the dimensions of the vessel, the relation of this speed to the nature of the gas and the conditions of the discharge will be comparatively simple, and probably much simpler than it is when the discharge is unsupported and the ions are travelling to the cathode through an abnormal cathode fall. For if the thermionic current is saturated (as it always was in these experiments) and the "space charge" neutralized, the distribution of potential throughout the vessel will be determined almost wholly by the geometrical form of the electrodes and of the walls. Further if collisions of the ions with molecules are rare, the speed of an ion reaching the grid will be simply that corresponding to the difference of potential between the grid and the place where the ion was formed. The distribution of these places will be almost independent of the nature and pressure of the gas, since, apart from small differences due to differences in the ionization potential, it will depend on the chance of an electron colliding with a molecule. Accordingly the mean energy of the ions reaching the grid should be independent of the nature and pressure of the gas and of the thermionic current.

In these circumstances, if the disintegration of the grid is determined by the number and energy of the positive ions reaching it, S should be independent of the pressure of the gas and the thermionic current; it should be a function of

V_0 and the nature of the gas. Further, measurements in different gases in the same vessel and for the same V_0 should be strictly comparable, for they will correspond to the same value of the mean energy. The only uncertainty in interpretation will arise from the fact that the relation between this mean energy and V_0, being determined by the geometrical arrangements of the electrodes, will not be known.

6. It was important to determine at the outset whether a pressure low enough to satisfy these conditions could be attained and yet the rate of disintegration remain large enough to be measurable. Examination of the characteristic of the discharge, and especially of the ratio of the grid to the anode current, suggested that the necessary conditions could be obtained; but the most conclusive evidence is obtained by measuring the disintegration with the same V_0 at different pressures and different thermionic currents.

Fig. 2.

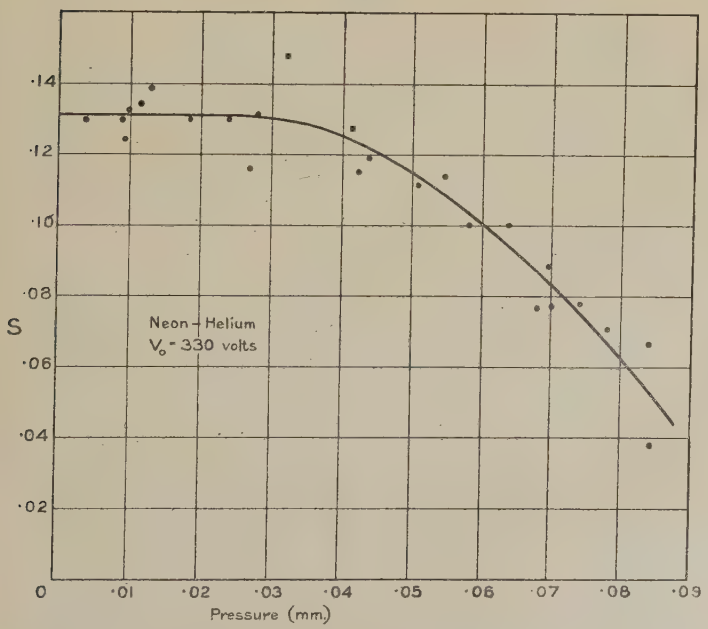


Fig. 2 shows the result of one series of measurements at different pressures, the gas being a mixture of atmospheric neon and helium containing about 25 per cent. of the latter.

It will be seen that at pressures below ·025 mm., S is independent of the pressure, while at higher pressures S decreases as the pressure increases. That result is completely in accordance with expectation, for at the higher pressures the ions will collide with molecules on their way to the grid, and the mean energy with which they arrive will be correspondingly reduced. Similar results were obtained in other gases; the pressure below which S becomes constant varies; it is greater for gases with long free paths (H_2, He) than for those with short free paths (Hg). All the measurements recorded below were made at pressures on the flat part of the curve. Again, many experiments were made to determine whether S was independent of the thermionic current. It was suspected that there might be some dependence because the ratio i_g/i_a decreased somewhat as i_a increased. This variation of i_g/i_a is doubtless due to recombination. In a vessel such as that used, with electrodes of area small compared with that of the walls, it is known * that some of the electrons follow very long paths, in the regions of weak field, so that the opportunities for recombination are considerable. This recombination might remove chiefly the ions which, if they arrived at the grid, would have little energy and cause little disintegration, so that an increase of S with i_a might have been anticipated. But no evidence of a dependence of S on i_g or i_a could be detected; the proportion of the ions of given speed disappearing by recombination must be approximately independent of that speed.

Lastly, it appeared that S was independent within wide limits of the potential of the cathode. Indeed V_a was of very little importance from any point of view so long as V_0 was constant; it could often be varied from 0·95 V_0 to 0·5 V_0 without a variation of more than a few per cent. in i_a or i_g. This independence of the discharge of the cathode potential is to be expected. For as long as the field at the cathode is sufficient to saturate, the thermionic emission i_a (which is always large compared with i_g) will be independent of V_a. Again, since the field is intense round the cathode, and V_a is large compared with the ionization potential, the volume within which the electrons have not yet gained the ionizing energy will be a small fraction of the whole volume of the vessel; consequently the number of ions formed per electron will not vary rapidly with V_0.

* Research Staff of G.E.C. Ltd., Phil. Mag. xlii. p. 230 (1921).

Results.

7. When these preliminary matters were settled and the method of investigation proved reliable, the factors determining disintegration could be investigated.

First it was shown that the impact of positive ions is necessary for disintegration. If V_g was negative, so that the grid acted as an anode, there was no disintegration whatever. Again when, as in the normal arrangement, V_g was positive, and the grid deflected from the cathode all but a very small proportion of the ions, very little disintegration of the cathode occurred. Definite measurements of the disintegration of the cathode were not made, but, by working with the grid sufficiently negative, it was possible to remove one-half of the material of the grid before the resistance of the cathode had increased by 1 per cent.

Second, it was shown that S is independent within wide limits of the temperature of the grid. Comparative experiments, made first with the grid relatively cold (it was always heated by the discharge to about 200° C.), and then with the grid heated by a battery to about 1200° C., disclosed not the smallest variation of the disintegration with the temperature. Higher temperatures were not used, because when thermionic emission from the grid begins, i_g is no longer a measure of the positive ions striking it. S was smaller than the normal for a newly inserted grid, but became normal as soon as surface impurities were removed.

Third, the rate of disintegration was determined for values of V_0 up to 700 volts in the following gases:—Hydrogen, helium, nitrogen, atmospheric neon-helium mixture, mercury vapour, argon. The results are shown in figs. 3 and 4. It was not possible to examine all these gases fully in a single vessel; for before sufficient observations have been made the disintegration of the grid has proceeded too far. But the observations in two vessels, which are shown in the figures, include between them all the gases mentioned. It is not to be expected that the absolute values of S in the two vessels will agree, because (see § 5) the same value of V_0 will correspond to different values of the mean energy, owing to slight differences in the distances between the electrodes. But the relations between the lines for the four gases which appear in both figures (Ar, Hg, Ne–He, He) are in general the same in each. Accordingly, the following conclusions seem justified:—

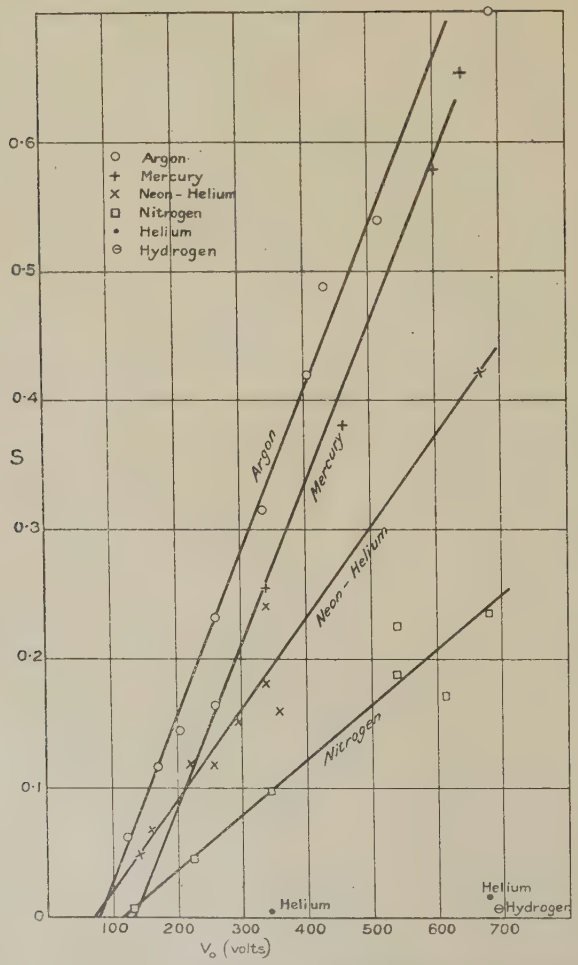
(*a*) The relation between S and V_0 is approximately of

the form

$$S = a(V_0 - V_0'), \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

where V_0' is not much greater than 100 volts. There was some tendency for the lines in figs. 3 and 4 to be concave to the horizontal axis, but as the observations were not

Fig. 3.



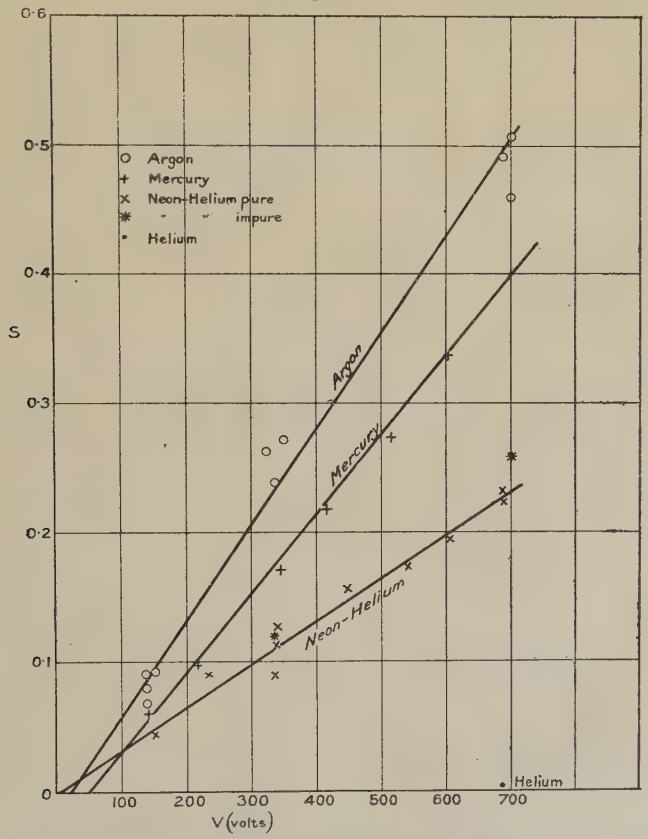
sufficiently accurate to determine the curvature the lines are drawn straight.

S may be greater than $\frac{1}{2}$. Since Cheney (*loc. cit.*), who presumably measured the current carried by charged atoms as well as by electrons, found the ratio of the negative

particles leaving to positive particles incident to be less than 0·1, it seems to follow that the metal particles are initially uncharged. In accordance with this view is the fact that there was no concentration of blackening round the anode.

(*b*) If the gases are placed in order of increasing magnitude for a given V_0, the series runs H_2, He, N_2, Ne–He, Hg, Ar, the difference between S for the two extreme numbers of this series being of the order of 100 to 1.

Fig. 4.



Attempts were made to measure S for potassium vapour. The discharge vessel was thoroughly evacuated and sealed off from the pump with a small quantity of the metal condensed on the walls; it was maintained in a small oven at about 100° C. in order that the vapour pressure of the metal might be sufficient to give a value of i_g of the same order as that given by the other gases. The experiments are difficult

and have not yet been carried to a successful conclusion. As far as they have gone they indicate that S for potassium vapour is of the same order as that for nitrogen—much more than for He, and much less than for Ar. This vapour was investigated because of the remarkable fact, discovered by the General Electric Company of America *, that a discharge giving a strong luminous glow with thermionic emission from a tungsten cathode can be maintained in sodium vapour (and apparently other alkali metal vapours) for many hours without serious disintegration of the tungsten.

Since they were using a pressure of the vapour of ·05 mm., the disintegration due to the discharge would be somewhat smaller than that calculated from S and the potential between the electrodes, because the mean free path would not be large compared with the distance between the electrodes, and the ratio of the mean to the maximum possible energy of the ions would be low. But our observations (if they are confirmed) agree with theirs in so far that S for metals (Hg, K) appears to be very much less than that for inactive elements of the same molecular weight.

8. In view of the remarkable decrease of disintegration in the independent discharge through the inactive gases produced by the addition of small traces of impurity (see p. 99), some experiments were undertaken on mixtures of argon, and of neon–helium with hydrogen and nitrogen. Quantitative measurements were difficult because the hydrogen and nitrogen disappear in the discharge much more rapidly than the inactive gases, so that after the discharge has lasted a short time the composition of the gas has changed. But the observations were sufficient to show that the effect found by Pintsch is not reproduced in the conditions of these experiments. (We do not mean to imply the smallest doubt that it is obtained in the conditions which he describes; indeed, we have been able to repeat all his observations.) Thus when 25 per cent. of H_2 was mixed with Ar, no certain decrease of S was produced; the absence of any decrease was doubtless due to the fact that Ar molecules were ionized, owing to their greater diameter, rather than the hydrogen. Similarly a mixture of 1 part (Ne–He) and 2 parts H_2 gave a value of S about half that of the (Ne–He) mixture. This mixture was known to contain at least 1 per cent. of impurity, probably N_2, and yet removal of this impurity by means of charcoal and liquid air produced no increase in S, but rather a decrease—which was doubtless due te the fact that the cooled charcoal removed more of the Ne than of the He. The points for the impure gas are shown on fig. 1.

* Brit. Patent 175032.

It was found that when a mixed gas was used, the ratio of S for the mixture to that for the pure constituents varied considerably with V_0. The explanation is doubtless that the proportion of the two constituents which are ionized varies with V_0. This observation explains why the line for the Ne–He mixture in figs. 2 and 3 cuts the lines for the pure gases.

9. The discharge which causes the disintegration is always accompanied by a decrease in the pressure of the gas; the gas disappears slowly, being absorbed in, or on the surface of, the metal deposited on the walls. This disappearance of the gas has been described by us in some detail in earlier papers. The experiments of this paper were not well adapted for studying the process, since the volume of gas in connexion with the discharge vessel was large and the measurement of pressure not very accurate. But some conclusions which are just worth record were reached.

The only gases which can be compared in this matter are argon, neon, helium, and nitrogen; for the disappearance of hydrogen is known to be anomalous, and that of the vapours could not be measured easily. Further, since the neon used was mixed with helium, conclusions concerning it are doubtful. It was found that the amount of gas disappearing per coulomb of positive ions reaching the grid was of the same order of magnitude whether the gas was neon–helium, helium, or nitrogen; a factor of 4 or less will cover all three gases. Nitrogen probably disappears most rapidly and helium least rapidly. The disappearance of argon is notably less than that of these three gases; indeed, it could never be established certainly that it disappeared at all. It follows that the gas disappearing per unit mass of metal deposited on the walls is far greater for helium than for the other gases, since the mass per coulomb is far less. It appears therefore that the rate of disappearance is determined by the rate of ionization of the gas rather than by the rate of deposition of metal; the disappearance cannot be explained simply as a trapping of the gas by the deposited metal. Probably the rate of deposition as important only in so far as it provides fresh surfaces to which the gas may adhere in the manner suggested by us in a previous paper *.

Conclusions.

10. The conclusion we would draw from these experiments is that the rate of disintegration S is a definite function of the nature of the positive ions and of their energy, increasing with the latter. There is no evidence that S is zero for any

* Research Staff of the G. E. C. Ltd., Phil. Mag. xliii. p. 914 (1922).

value of V_0 less than the ionization potential—for the prolongation of the lines to cut the axis cannot be relied on—and there is no evidence of any exceptional change in S when V_0 is in the neighbourhood of the normal cathode fall, which varies from about 140 volts for Ne to 330 for H_2. Further, there is no immediately obvious connexion between S for a given V_0 and the other properties of the ion. In general S increases with the mass of the ion, as other workers have found, but the order of Ar and Hg, and (probably) of Ne and K are exceptions. Again, the inactive gases appear especially active in disintegration, but here He is an exception. On the other hand, it appears that the metals may be especially inactive. These relations are very puzzling, and as yet we offer no theory to explain them; the next step in developing an adequate theory of cathode disintegration must be to correlate the disintegrating power of an ion with some other of its properties.

But it is already possible to see that information derived from such experiments as these will throw light upon the problems of the unsupported discharge. Thus, in that form of discharge, the loss of mass by disintegration per coulomb carried between the cathode and the anode is proportional to $V-V'$, where V is the potential between the electrodes, and V' of the order of the normal cathode fall; *e. g.* Kohlschütter gives $V'=495$ volts for air. There is practically no sputtering so long as the cathode fall is normal. Interpreted in the light of our experiments, this must mean that with a normal cathode fall, either the current at the cathode is conveyed almost wholly by electrons, or, more probably, that the positive ions arrive at the cathode with energies far less than that corresponding to the normal cathode fall. It is only when the cathode fall is abnormal that the positive ions strike it with any considerable velocity.

Again, the reduction of the disintegration with abnormal cathode fall in the inactive gases due to the presence of small quantities of impurity must mean either that the energy of the ions arriving at the cathode is reduced or that ions of the impurity, rather than of the inactive gas, carry the current to the cathode. Since mercury is as effective as other impurities, the first alternative seems the more probable.

Generally, measurements of disintegration will enable a distinction to be made between current conveyed by positive ions arriving at an electrode and by electrons leaving it. In the hope of applying the knowledge gained to such problems, we propose to carry these problems further, and, in particular, to get a more definite relation between S and the energy of the ions.

VIII. *On the Molecular Scattering of Light in Dense Vapours and Gases.* *By* C. V. Raman, *M.A., Hon. D.Sc., Palit Professor of Physics in the Calcutta University*; *and* K. R. Ramanathan, *M.A., Demonstrator in Physics, Maharajah's College, Trivandrum, and University of Madras Research Scholar* *.

1. *Introduction.*

As is well-known, the theory advanced by the late Lord Rayleigh † explaining the blue light of the sky as due to the molecular scattering of sunlight has of recent years been amply confirmed by accurate measurements of sky-radiation and of the attenuation of the sun's rays in traversing the earth's atmosphere, and also by the laboratory experiments of Cabannes ‡ and of the present Lord Rayleigh § which have demonstrated the scattering of light by dust-free gases and shown that its magnitude is in agreement with the theory. Apart from its special application to the explanation of the blue light of the sky, the subject of the molecular scattering of light possesses very great interest, standing as it does in intimate relationship with the fundamental problems of optics, such as the propagation of light and the reflexion, refraction, dispersion, and extinction of light in transparent media. Its study has already yielded information which is of first-rate importance in testing theories of atomic and molecular structure, and promises to prove even more fruitful in future in the same direction. It is also not improbable that the study of the problems of molecular scattering of light may serve to elucidate the nature of radiation itself, that is, to determine whether it consists of some form of continuous disturbance moving through space, as contemplated by Maxwell's electromagnetic theory, or whether it consists of discrete entities or "quanta," as conceived of by Einstein. It is proposed in this paper to consider the very important problem of the scattering of light in *dense* media, such as vapours and gases under high pressure, in which circumstances, as was pointed out by the late Lord Rayleigh himself, the simple treatment proposed by him is not applicable. Fresh experimental evidence bearing on the subject will also be referred to.

* Communicated by the Authors.

† Phil. Mag. xlvii. pp. 375–384 (1899). Scientific Papers, vol. iv. p. 397.

‡ *Annales de Physique*, xv. pp. 1–150.

§ Proc. Roy. Soc. xciv. p. 453 (1918) *et seq.*

2. *Rayleigh's Theory of Light-Scattering.*

As a preliminary to the discussion of the problem of light-scattering in highly compressed gases and vapours, it is useful to recall very briefly the principles on which the case of gases under ordinary conditions was dealt with by Lord Rayleigh. Reduced to its essentials, the treatment is on the following lines:—

The individual molecules in a gas through which the primary waves of light pass are regarded as secondary sources of radiation, each molecule acting more or less as it would in the absence of its neighbours. There is, of course, a definite phase-relation between the primary wave when it reaches a given molecule and the secondary wave emitted from it. *In the direction of propagation of the primary waves* the secondary radiations emitted by all the molecules in a given layer are in identical phase, for the differences in the phase of the primary wave when it reaches different molecules are exactly compensated by the acceleration or retardation due to the scattered waves having to traverse a shorter or greater path, as the case may be. *In other directions*, however, owing to the molecules being distributed at random within the volume of the gas, the phases of the scattered waves do not stand in any invariable relation to each other, and hence, in order to find the average expectation of intensity of the scattered light emerging from within the gas, the *intensities*, not the amplitudes of the waves scattered by the individual molecules, should be added up. In all ordinary cases there is very little difference of phase between the primary wave reaching an individual molecule and the secondary wave sent out by it so far as the direction of original propagation of the wave is concerned. When, however, the effect of all the molecules contained in a stratum parallel to the plane of the primary wave is integrated by the usual method of subdivision of the stratum into Fresnel zones, and the resultant is combined with the primary wave, a change of phase appears which may be identified as the retardation associated with the passage of waves through a refractive medium. A relation is thus obtained between the scattering power of the molecules, their number per unit volume, and the refractivity of the medium. Thus, taking the light-vector in the primary waves to be represented by $R_0 \cos(\omega t - lx)$, where x is measured from the position of the scattering molecule, the vector in the scattered wave arising from it may be written as

$$A_1 \cos(\omega t - lr) R_0 \sin\theta / r, \quad . \quad . \quad . \quad . \quad (1)$$

when r is the distance from the molecule and θ is the angle between the vibration at the origin and the scattered ray. On carrying out the calculations indicated, the relation obtained is

$$n_1 A_1^2 = \frac{4\pi^2(\mu-1)^2}{n_1\lambda^4}, \quad \ldots \ldots \quad (2)$$

where n_1 is the number of molecules in unit volume, λ is the wave-length of the light, and μ is the refractive index of the gas.

The energy scattered by the molecules in the interior of the gas must be derived from the primary beam, and hence the intensity of the latter must suffer an attenuation as it passes through the medium. The coefficient of attenuation may be readily evaluated by a simple calculation of the total energy scattered by an individual molecule, and then multiplying this by the number n_1 of molecules per unit volume. We thus obtain the intensity of the transmitted light to be $I = I_0 e^{-kx}$, where k the coefficient of attenuation is given by the relation

$$k = \frac{32\pi^3(\mu-1)^2}{3n_1\lambda^4}. \quad \ldots \ldots \quad (3)$$

This expression for the attenuation coefficient may also be derived directly by a more accurate investigation, which does not neglect the small difference of phase between the primary wave and the secondary waves originating at a molecule. It is then found that on compounding the effects of the primary wave with those of the secondary wave arising from a stratum of molecules, there appears, in addition to the alteration of phase of the primary wave, also a small diminution in its intensity which is exactly that expected in view of the lateral scattering of part of the energy.

3. *Justification of the Principle of Random Phase.*

It will be seen from the foregoing that the Rayleigh law of scattering depends for its validity on the principle that the phases of the waves laterally scattered by the individual molecules are absolutely at random. This point is clearly brought out in the detailed analyses of the problem which have been given H. A. Lorentz * and by the late Lord Rayleigh himself †, and it appears that the necessary random character of the phases of the scattered waves can only arise

* Proc. Roy. Soc. of Sciences, Amsterdam, xiii. p. 92 (1910).
† Phil. Mag. xxxvi. pp. 429-449 (1918).

from a random arrangement in space of the molecules under consideration. If we deny the random character of this distribution of phase of the scattered waves, we must also deny the applicability of Rayleigh's law expressing the intensity of the scattered light in terms of the refractivity of the medium. The fact that even in a gas at atmospheric pressure there are about 10^6 molecules in a cube of the size of a wave-length of light, so that the scattered waves arising from adjacent molecules differ in phase by only 10^{-2} of the period, has been felt as a difficulty by some writers in this connexion. In fact, Larmor * has raised the question whether in view of this closeness of packing we are really justified in considering the phases of the individual scattered waves to be distributed entirely at random. It may be easily shown, however, that no appreciable deviation from a random distribution of phase, and therefore, also, from the Rayleigh law of scattering, can arise so long as the pressure-volume relation of the gas does not sensibly deviate from Boyle's law. For, all that is necessary is that the chance that a given molecule finds itself in a specified finite element of volume should be a random one, that is, is sensibly uninfluenced by the other molecules in the gas. This is true provided the total volume occupied by all the molecules, or rather by all their spheres of mutual action, forms a small fraction of the total space in the containing vessel, which is obviously also the condition that the compressibility of the gas should not appreciably deviate from that given by Boyle's law.

Larmor has made the suggestion † that Rayleigh's principle of random phase is not correct in the form in which it is stated, but that the random character of the phases of the scattered waves when they reach the observer is secured by the Doppler effect, due to the motion of the scattering molecules. The view he proposes is, in fact, that the phases of the scattered radiations emerging from the column of gas may be correlated at first, but after traversing 10^6 wave-lengths, or 50 cms., they would have become fortuitous and the energy-effects thus additive. Larmor's suggestion is, however, not acceptable. For, if it were correct, we should find that the aggregate intensity of the scattered light should be small immediately after emergence from the column of gas, that is, when it is observed within a distance of a centimetre or two from the track of the primary beam, and should increase at a greater distance from it. Such a result is

* Phil. Mag. xxxvii. p. 161 (1919).

† *Loc. cit.*

obviously quite inadmissible, besides being contrary to experience. The fallacy lies in Larmor's assumption that the change of wave-length (Doppler effect) has an effect on the relative phases of the scattered waves, whereas in reality it has none. To make this clear, we may consider two neighbouring molecules A and B. The scattered waves originating from them travel outwards with an identical velocity which is quite unaffected by any movements of these molecules. The phase-difference at any epoch therefore remains unaffected as the waves move out, being exactly the same as when the portions of the wave-train under consideration left the molecules. In other words, the phase-difference at each stage is exactly the same as if the molecules had remained fixed from the instant of emission of the scattered light. The scattering from any appreciable volume of gas would thus remain unaffected if all the molecules were assumed suddenly to be fixed in their instantaneous positions, and the Doppler effect due to their movement exerts no influence whatever on the observed results. Larmor's suggestion, therefore, clearly fails, and we have to fall back upon the view already expressed that the *principle of random phase of the scattered waves rests upon the random character of the distribution of the molecules in space in a gas obeying Boyle's law.*

4. *Application of Boltzmann's Distribution Law.*

It is clear from the preceding discussion that if the distribution of the molecules of a gas or vapour in space is *not* a random arrangement, the intensity of the light scattered would not be given by the Rayleigh law, that is, would cease to be simply proportional to the number of molecules per unit volume. In the usual treatments of the kinetic theory of gases, the molecules are conceived to exert forces on each other, and if the time during which the forces acting on any individual molecule are sensible forms an appreciable fraction of the whole time, the arrangement of the molecules would not be a random one. The influence of the molecules on each other would not then be negligible, and the energy-effects of the individual molecules could not be added up to find the total light scattered in the manner contemplated by the simple theory. It is, however, possible to make a calculation of the altered scattering power of the gas. Let $\phi(r)$ be the law of repulsive force between any two molecules whose centres are at a distance r from each other. By a suitable assumption regarding the nature of $\phi(r)$, it is possible to

include in this function the cohesive or van der Waals's forces which act when the molecules are not in collision with each other, and also the large repulsive forces occurring during actual impact. According to a well-known theorem due to Boltzmann*, the chance that two molecules find themselves at a finite distance r apart is less than the chance that they are at a very great distance apart in the ratio

$$e^{-\frac{1}{\mathrm{RT}}\int_r^\infty \phi(r)dr}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where $\int_r^\infty \phi(r)dr$ obviously expresses the potential energy of the configuration. The probability of groups of three, four, or more molecules being found at specified distances apart may be similarly expressed in terms of the potential energies of the respective configurations. If we assume that the deviation from Boyle's law is not very large, expression (4) will suffice to enable us to take into account the mutual influence of the molecules in determining their ordering in space. Consider a sphere of radius l drawn round the instantaneous position of a given molecule as centre. If n be the whole (actual) number of molecules in the volume of the gas or vapour, the number to be expected within this sphere of volume $\frac{4}{3}\pi l^3$ is

$$\int_0^l \mathrm{A} \cdot 4\pi r^2 dr\, e^{-\frac{1}{\mathrm{RT}}\int_r^\infty \phi(r)dr} \quad . \quad (5)$$

where A is a constant to be presently determined. Integrating by parts, (5) reduces to

$$\mathrm{A}\left[\frac{4}{3}\pi r^3 e^{-\frac{1}{\mathrm{RT}}\int_r^\infty \phi(r)dr} - \frac{1}{\mathrm{RT}}\int_0^l \frac{4}{3}\pi r^3 \phi(r)dr\, e^{-\frac{1}{\mathrm{RT}}\int_r^\infty \phi(r)dr}\right]_0^l . \quad (6)$$

To enable this expression to be reduced, we will assume that the force $\phi(r)$ between two molecules vanishes sensibly when $r \gtrless \sigma$, where σ is the range of molecular action. The expression (6) then reduces to

$$\mathrm{A} \cdot \frac{4}{3}\pi l^3 - \frac{\mathrm{A}}{3\mathrm{RT}}\int_0^\sigma 4\pi r^2 r\, \phi(r)\, dr\, e^{-\frac{1}{\mathrm{RT}}\int_r^\infty \phi(r)dr}, \quad . \quad (7)$$

provided $l \gtrless \sigma$. When l is very great, the first term in (7) will be very large in comparison with the second, and since (7) must then give the total number of molecules, we find at once $\mathrm{A} = \frac{n}{v}$. We may denote the integral in (7) by C;

* *Vorlesungen über Gastheorie*, II. See also Jeans' 'Dynamical Theory of Gases,' 3rd ed. p. 132.

and putting $l=\sigma$, we find that, on the average, the number of molecules which we may expect to find within the sphere of action of the specified molecule is

$$\frac{n}{v}\cdot\frac{4}{3}\pi\sigma^3-\frac{nC}{v.3RT}, \quad . \quad . \quad . \quad . \quad . \quad (8)$$

and is thus smaller than the number corresponding to a random distribution of the molecules by $\frac{nC}{v.3RT}$. (If C be negative, the number will be greater.)

In order to evaluate the intensity of the scattered light, we may now apply the general method developed by the late Lord Rayleigh *. The disturbance due to the scattered waves may be represented by the summation

$$\Sigma\cos(pt+\epsilon_1), \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

where ϵ_1 represents the phase of the scattered wave arising from a molecule and is easily expressed in terms of the position of the molecule and the directions of the incident and scattered rays. The *intensity* of the scattered light is obviously represented by

$$\begin{aligned}&[\cos\epsilon_1+\cos\epsilon_2+\cos\epsilon_3+\ldots.]^2\\&+[\sin\epsilon_1+\sin\epsilon_2+\sin\epsilon_3+\ldots.]^2\\&=n+2\Sigma\cos(\epsilon_1-\epsilon_2), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)\end{aligned}$$

where n is the number of molecules and the summation $\Sigma\cos(\epsilon_1-\epsilon_2)$ has to be made for all the $\frac{1}{2}n(n-1)$ pairs of molecules contained in the volume. For a random distribution, $\Sigma\cos(\epsilon_1-\epsilon_2)$ vanishes, as has been shown by Lorentz and by Rayleigh. In the present case, the distribution differs from a random one in that the average number of molecules present within the sphere of action of a molecule is smaller by $\frac{n}{v}\frac{C}{3RT}$. The summation indicated in (10) will not therefore reduce to n simply. Since the radius of action σ of the molecules is much less than the wave-length of the light, the value of $(\epsilon_1-\epsilon_2)$ for the missing molecules may without appreciable error be taken to be zero, and hence, when the summation is carried out for all the molecules in the volume v, (10) will reduce to

$$n\left(1-\frac{n}{v}\cdot\frac{C}{3RT}\right). \quad . \quad . \quad . \quad . \quad (11)$$

* Phil. Mag. xxxvi. pp. 429–449 (1918).

From this it is seen that if C be positive, the intensity of the scattered light is less than that given by the Rayleigh law, while if C be negative, it is *greater*.

In order to interpret the result contained in (11), we notice that the integral in the second term of (7) which we have denoted by C when multiplied by $\frac{1}{2}\frac{n^2}{v^2}$ represents the *virial function* for the gas or vapour. The virial equation is in fact simply*

$$pv = n\mathrm{RT} + \frac{1}{6}\frac{n^2}{v}\mathrm{C}. \quad . \quad . \quad . \quad . \quad . \quad (12)$$

By differentiating (12), the compressibility β of the gas or vapour may be written as

$$\beta = -\frac{1}{v}\frac{dv}{dp} = \frac{1}{n\frac{\mathrm{RT}}{v}\left(1+\frac{n}{v}\frac{\mathrm{C}}{3\mathrm{RT}}\right)} \quad . \quad . \quad . \quad (13)$$

$$= \frac{v}{n\mathrm{RT}}\left(1-\frac{n}{v}\frac{\mathrm{C}}{3\mathrm{RT}}\right), \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

to the same degree of approximation as used previously. On comparing (14) with (11), it is seen that the intensity of the light scattered per unit of volume, instead of being proportional to $\frac{n}{v}$ as required by the Rayleigh law, is proportional to

$$\frac{n^2}{v^2}\mathrm{RT}\beta, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

and thus varies directly as the absolute temperature, the square of the density of the substance and its compressibility. In the case of a gas obeying Boyle's law, $\beta = \frac{1}{p}$ and $pv = n\mathrm{RT}$, and (15) becomes simply $\frac{n}{v}$ in agreement with the Rayleigh law.

5. *The Theory of Density Fluctuations.*

The result deduced by us in the preceding section from the principles of the kinetic theory for the case of moderately compressed gases agrees with the general formula obtained

* This agrees with Keesom's approximate equation of state, containing only the first two terms in the expansion of the virial function. Proc. Roy. Acad. Sciences, Amsterdam, vol. xv. p. 264.

in an entirely different manner by Einstein * in developing an explanation given by Smoluchowski † of the phenomenon of opalescence exhibited by fluids in the vicinity of the critical state. Einstein's reasoning is of a very general character and makes no use whatever of the molecular hypothesis. It is based on the following train of ideas. If, in the condition of equilibrium, the medium is conceived of as a completely uniform homogeneous system, there can be no scattering of the light-waves passing through it. Actually, however, as Smoluchowski showed from thermodynamical considerations, the density of the substance must vary from point to point, and also fluctuate from instant to instant, the density as measured by our ordinary appliances being merely a statistical average. Hence, in framing the electromagnetic equations, the dielectric constant (which is assumed to vary with the density in the manner contemplated by the Lorentz-Mossotti formula) must be regarded as varying from point to point, and the solution of the equations for wave-propagation contains, besides the regularly transmitted waves, also an opalescence radiation which is a volume-effect and represents the scattered light resulting from the local inequalities of density. Einstein's treatment naturally therefore falls into two parts: firstly, the determination of the magnitude of the fluctuations of the dielectric constant for the elements of volume, secondly, the calculation of the opalescent radiation.

The magnitude of the density-fluctuations is calculated by Einstein in the following manner:—Boltzmann's well-known thermodynamical principle connects the entropy S of a system with the probability W to which the entropy S relates by the equation

$$S = R \log W + \text{constant}, \quad . \quad . \quad . \quad . \quad (16)$$

where R is the gas-constant. Einstein points out that the entropy S of the whole system does not completely define the physical conditions of the small individual parts of the system at any instant, and hence (16) should be understood as referring to an *ensemble* of different possible conditions of the system in each of which its entropy is S. If we consider particular parts of the system and particular instants of time, fluctuations from the ideal thermodynamic equilibrium must appear, the magnitude of which is determined by the corresponding changes of entropy of the parts of the system. For an isothermal compression or rarefaction, the change of

* *Annalen der Physik*, xxxiii. p. 1294 (1910).
† *Loc. cit.* xxv. p. 205 (1908).

entropy is simply the work done divided by the absolute temperature. Boltzmann's theorem may be written in the form

$$W = e^{\frac{S-S_0}{R}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

where S_0 is the entropy in the state of thermodynamic equilibrium. Taking ν to be the deviation of some observable physical variable, *e. g.* in this case the density, from the value corresponding to thermodynamic equilibrium, Einstein writes (17) in the form

$$W(\nu)d\nu = Ae^{-\frac{1}{2RT}a\nu^2} d\nu, \quad . \quad . \quad . \quad . \quad (18)$$

where $W(\nu)d\nu$ is the probability that ν lies between ν and $\nu + d\nu$, A is a constant, and $\frac{1}{2}a\nu^2$ is the work done in altering ν from 0 to ν. (18) evidently has the same form as Gauss's error-law, and it follows that

$$\overline{\frac{1}{2}a\nu^2} = \frac{1}{2}RT, \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

where the bar indicates the *mean* value of the quantity under consideration. If we identify ν with $(\Delta\rho)$ the fluctuation of density, its mean square may be evaluated from (19) and written in the form

$$\overline{(\Delta\rho)^2} = \frac{RT\beta\rho_0^2}{\phi}, \quad . \quad . \quad . \quad . \quad . \quad (20)$$

where β is the isothermal compressibility, ϕ is the small element of volume under consideration, and ρ_0 is the mean density.

For the deduction of the light-scattering from (21) with the aid of Maxwell's electromagnetic equations, reference should be made to Einstein's original paper. We may follow here an abbreviated treatment which gives an identical result.

At right angles to the incident light the intensity of scattered light is given by the expression *

$$\frac{\pi^2\phi^2(\Delta\epsilon)^2}{2\lambda^4}\frac{1}{r^2}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

$(\Delta\epsilon)$ being the variation of the dielectric constant.

Now, as the individual local variations of the dielectric constant are irregularly distributed, the phases of the various scattered beams are also quite arbitrary, and hence for calculating the total intensity of scattered light we have

* Rayleigh, Phil. Mag. p. 81 (1881); Sc. Papers, vol. i. p. 526.

merely to sum up the above expression over the total volume v. The expression contains a factor

$$\Sigma(\Delta\epsilon)^2\phi^2,$$

which may be evaluated by use of the Mosotti-Lorentz law

$$\frac{\epsilon-1}{\epsilon+2}=\text{const.}\,\rho.$$

Differentiating, we obtain

$$(\Delta\epsilon)^2=\frac{(\epsilon-1)^2(\epsilon+2)^2}{9}\left(\frac{\Delta\rho}{\rho_0}\right)^2. \quad . \quad . \quad . \quad (22)$$

Substituting in this the value given above for fluctuation of density, we obtain

$$\frac{(\epsilon-1)^2(\epsilon+2)^2}{9}\mathrm{RT}\beta\Sigma\phi=\mathrm{RT}\beta\frac{(\epsilon-1)^2(\epsilon+2)^2}{9}.v. \quad (23)$$

Thus the intensity of light scattered by a cubic centimetre of fluid at right angles to the incident rays is

$$\frac{\pi^2}{18}\mathrm{RT}\beta\frac{(\epsilon-1)^2(\epsilon+2)^2}{\lambda^4}=\frac{\pi^2}{18}\mathrm{RT}\beta\frac{(\mu^2-1)^2(\mu^2+2)^2}{\lambda^4}. \quad (24)$$

In the cases of gases and vapours, whether obeying Boyle's law or not, the formula (24) may be appreciably simplified by assuming $(\mu^2+2)=3$, and $(\mu^2-1)^2=4(\mu-1)^2$ approximately, and it may then be written in the form

$$\frac{2\pi^2(\mu-1)^2\mathrm{RT}\beta}{\lambda^4}. \quad . \quad . \quad . \quad . \quad . \quad (25)$$

If the gas or vapour obeys Boyle's law, $\beta=\frac{1}{p}$ and $pv=n\mathrm{RT}$, and (25) becomes identical with the Rayleigh law. In other cases, (25) is seen to agree with formula (15) when the omitted multiplying factors are restored.

It may be pointed out that the expression giving the magnitude of the density-fluctuation might have been derived immediately from the principle of equipartition of energy, as is indeed evident from the form of the relation (19). In the case of a gas obeying Boyle's law in which the molecules are distributed at random throughout the space, the mean square of the density-fluctuation may also be derived by direct application of the theory of probabilities, and the intensity of the scattering calculated therefrom. This has been done by Lorentz *, and the result naturally agrees with that obtained by Rayleigh.

* Proc. Roy. Soc. Amsterdam, xiii. p. 92 (1911).

6. *Experimental Tests of the Einstein-Smoluchowski Formula.*

The expression given by Einstein for the light-scattering in a direction transverse to the primary beam, viz. :

$$\frac{\pi^2}{18}\frac{R_1 T\beta}{N_1}\frac{(\mu^2-1)^2(\mu^2+2)^2}{\lambda^4},$$

where R_1 and N_1 represent the gas-constant and the Avogadro-number respectively for a a *gram-molecule*, was originally developed by him to explain the opalescence exhibited by fluids in the vicinity of the critical state. It should, however, possess a far wider and *general* application, and indeed be regarded as superseding the Rayleigh law of scattering, which should only be valid in the special case of gases accurately obeying Boyle's law. It is thus of great importance to test the validity of the Einstein-Smoluchowski formula over a very wide range of cases. So far as the authors are aware, the only experimental tests that have so far been made of it for the case of a single substance in the gaseous state are those of Kammerlingh Onnes and Keesom * on the opalescence of ethylene vapour, but the work of these investigators was confined to a range of only 2·5 degrees of temperature *above* the critical temperature of this substance. A far more comprehensive and general test of the formula for several substances over the widest possible range of temperatures and pressures is obviously desirable, and researches with this object have been undertaken in the authors' laboratory. The first substance chosen was ether. This substance has a high refractivity and therefore a large scattering power in the state of vapour, and its critical temperature and pressure are moderate, and it also possesses a high degree of chemical stability under the action of heat and light. It is therefore very well suited for quantitative work. The detailed account of the determinations carried out with it by one of us (K. R. Ramanathan) will be published in a separate paper, and it will suffice here to indicate the general scope of the work and the results obtained. The light-scattering power of ether has been determined separately for the substance in the saturated vapour and in the liquid phases for temperatures from 30° C. to the critical temperature 194° C., and for the gaseous phase from 194° C. to 217° C. . The vapour and gaseous phases deviate very largely from Boyle's law, the compressibility being larger than the reciprocal of the pressure, especially at the higher temperatures, and the

* *Loc. cit.* (1908).

measurements throughout the whole range show that the *Rayleigh law of scattering completely breaks down* and strikingly confirm the correctness of the Einstein-Smoluchowski formula. The scattering power of the vapour when it is dense is in fact several times larger than that given by the Rayleigh law, and this feature is more and more pronounced as the critical temperature is approached from either side. The Einstein-Smoluchowski formula also shows itself to be equally applicable to the case of liquids over the whole range of temperature tried.

7. *Polarization of the Scattered Light.*

It should be noted that, according to Einstein's investigation, the light scattered in a direction transverse to the primary beam should be completely polarized. Actually, however, this is not the case. The imperfectness of the polarization of the light scattered by gases was first discovered by Strutt (the present Lord Rayleigh), and measurements of the magnitude of the effect have been made by a photographic method by him * and by Cabannes † and also by Gans ‡. The imperfect polarization can be detected visually even in the case of dust-free air at atmospheric pressure §, and with the gases at high pressure, the increased intensity enables *visual measurements* to be carried out. In the case of ether vapour, the effect is very feeble, the polarization in the transversely scattered light being nearly complete (about 98·5 per cent.), but the great intensity of the scattered light enables *visual* measurements to be carried out, and it is found that with increasing density of the vapour, the percentage of polarization does not appreciably change. In the case of the liquid, specially interesting effects are observed which will be dealt with in a separate paper.

Cabannes ‖ has shown that owing to the imperfect polarization, the intensity of the transversely scattered light is increased and is greater than that given by the Rayleigh law in the ratio $\frac{6(1+\rho)}{6-7\rho}$, where ρ is the ratio of the weaker to the stronger component of polarization. It is found that the Einstein-Smoluchowski formula should be modified by a similar correcting factor in order to give results in agreement with experiment. The correction becomes very important when ρ is large.

* Proc. Roy. Soc. xciv. p. 453 (1918).
† *Ann. de Physique*, xv. pp. 1–150.
‡ *Ann. der Physik*, lxv. p. 97 (1921).
§ 'Nature,' Jan. 19, 1922.
‖ *Loc. cit.*

8. *Relation of Molecular Scattering to Critical Opalescence.*

As we have already seen, the Einstein-Smoluchowski formula shows that there is a perfect continuity between the comparatively feeble light-scattering shown by gases and vapours under ordinary conditions and the extremely large scattering or opalescence which develops as the critical state is approached. At the critical point itself, the compressibility of the substance is infinite, and it follows that the Einstein-Smoluchowski formula gives an infinite value for the scattering power. This is obviously inadmissible, and Einstein * himself has pointed out that for a *very minute* region at and round the critical point, the neglect of second-order terms in the expression for the work done in altering the density is not permissible and the formula requires modification. This point has, however, not much importance for our present purpose, as the range of temperature over which the neglect of the second-order terms is not permissible is exceedingly small. It may be mentioned that the matter has been further discussed by Ornstein and Zernike †, who have proposed an amended formula applicable in the immediate neighbourhood of the critical point which leads to a finite value for the scattering at this point. The bearing of Ornstein and Zernike's results on the phenomena observed with ether will be dealt with separately.

9. *The Case of Carbon Dioxide and the permanent gases.*

The experimental confirmation of the Einstein-Smoluchowski formula for the case of ether vapour, and also in the case of various ordinary liquids, shows that its theoretical basis is thoroughly firm and that the formula may be justly expected to apply equally well in the case of other gases and vapours. It is therefore very important that the case of carbon dioxide and the permanent gases, *e. g.* oxygen, nitrogen, argon, &c., should be thoroughly investigated. The scattering power of saturated carbon dioxide vapour at 21° C. (that is, ten degrees below the critical temperature) has been determined by the present Lord Rayleigh, and found to be 102 times that of the gas at atmospheric pressure. The compressibility of carbon dioxide at this temperature is very great, and the Einstein-Smoluchowski formula shows that the scattering power of saturated carbon dioxide at 21° C. should have been about 855 times that of the gas at atmospheric pressure. The observed result thus apparently supports the Rayleigh law of scattering (the density of carbon dioxide at this temperature being about

* *Ann. der Physik,* xxxiii. p. 1275 (1910).

† Proc. Amsterdam, xvii. p. 793 (1914).

110 times the density of atmospheric pressures), and not the Einstein-Smoluchowski formula. The experimental result is however open to question, as Rayleigh mentions that his observation tube had a leak, and the density of the gas was therefore probably much less than that corresponding to the saturation pressure. Some observations made in a Cailletet tube containing CO_2 liquid and vapour by the authors show that both carbon dioxide liquid and vapour scatter light very strongly at and for a few degrees below the critical temperature, but the scattering power falls off very quickly with decreasing temperature. It is hoped to carry out a series of measurements at different temperatures with improved apparatus, and to obtain data in order decisively to settle the question whether the Einstein-Smoluchowski formula does really break down for carbon dioxide. Probably standing in very close relation to this is the fact that has been observed by Kammerlingh Onnes and his co-workers at the Leiden Laboratory, that pure oxygen, nitrogen *, and argon do not show the phenomenon of critical opalescence. It is obviously of the greatest importance to investigate whether in these cases the Einstein-Smoluchowski formula is applicable or not.

Till further data are available, it is probably not safe to venture an opinion, but it seems clear that the applicability of the thermodynamic reasoning of Einstein and Smoluchowski stands on an unchallengeable basis at least at ordinary temperatures and so far as gases and liquids are concerned. It will be noticed that the formula was derived on the basis of the classical theory of light in which it is assumed to consist of some form of continuous wave-propagation. In view of recent work in various modern fields of physics, and the known failure of the classical theory in explaining the facts regarding the emission and absorption of light, the question arises whether we are justified in assuming that the classical wave theory is necessarily capable of giving a strict quantitative explanation of the facts regarding the molecular scattering of light? Is it not possible that in the case of substances whose molecules have a relatively small light-scattering power, the infinite sub-divisibility of the energy of the incident waves which the classical theory assumes might break down? Then again is it not possible that the statistical-mechanical considerations on which the validity of Boltzmann's principle rests might require modification in the sense of the quantum theory, particularly in dealing with fluids at low temperatures? The answer to these questions can only be given when a more thorough experimental investigation of the facts of molecular

* Proc. Amsterdam Academy, xviii. p. 950 (1915).

scattering in dense vapours has been carried out. Such investigation has been undertaken by the authors, and it is hoped that before long the necessary data will be available.

10. *Summary.*

The paper considers the problem of the molecular scattering of light in dense vapours and gases *not* obeying Boyle's law. The following are the principal results :—

(1) It is shown by a discussion that the Rayleigh law of scattering, which is based on the assumption of a random distribution of the molecules in space, should not be applicable in these cases.

(2) Assuming the non-random distribution of the molecules in space to be given by Boltzmann's theorem, a calculation is made of the scattering power of the medium. The result obtained is interpreted with the aid of the virial equation, and it is shown that the scattering power per unit volume is proportional to the absolute temperature, to the square of the density of the substance, and to its compressibility. If the gas obeys Boyle's law, the expression reduces to the Rayleigh law.

(3) The result thus deduced directly from the principles of the kinetic theory agrees with the general formula derived from the thermodynamical theory of density fluctuations, developed by Einstein and Smoluchowski, of which a brief account is given.

(4) Experiments carried out with ether vapour and liquid over a large range of temperature both below and above the critical point completely confirm the Einstein-Smoluchowski formula for this substance.

(5) The corrections of the formula necessary in view of the imperfectness of the polarization of the transversely scattered light, and also in the immediate neighbourhood of the critical point when the compressibility becomes excessively great, are indicated.

(6) The cases of carbon dioxide and of the permanent gases are referred to and the possible significance of the results obtained in their cases is discussed.

Calcutta, India.
April 22nd, 1922.

Note added in proof (*6th August*) : Since the paper was sent in, experiments on the light-scattering in the vapours of benzene and pentane at different temperatures and also in *carbon dioxide* at high pressure contained in sealed bulbs have been carried out, and these definitely confirm the Einstein-Smoluchowski formula and show that the Rayleigh law of scattering is not valid except in gases obeying Boyle's law.

IX. *On the Anomaly of Strong Electrolytes with special Reference to the Theories of J. C. Ghosh.* *By* Henry J. S. Sand, *D.Sc., Ph.D., F.I.C.; Sir John Cass Technical Institute, London* *.

Synopsis:—Part I. Introduction. — Application of Boltzmann's theorem to obtain an estimate of the degree of association of a completely ionized electrolyte immersed in a medium of uniform dielectric constant.— Application of Born & Landé's formula to obtain an estimate of the probability that an ion obeys the inverse square law.—Ghosh's virial formula.—Thermodynamical deduction of an equation for the heat of dilution of an electrolyte whose osmotic pressure obeys a law such as that of Ghosh or Milner.—Test of Ghosh's and of Milner's osmotic pressure formulæ by experimental results. Part II. Ghosh's theory of electrolytic conductivity.—Adjustment of the subsidiary hypotheses so that they yield Ghosh's conductivity formula.—Discussion of constancy of conductivity under varying electrical conditions in its relation to a theory of complete dissociation with uniform dielectric constant of the medium.—Rejection of the hypothesis of uniform dielectric medium, reference to similar conclusion by Milner.—Discussion of the rôle of molecules of solvent and of solvated ions.—Test of Ghosh's conductivity formula by experimental results, effect of substituting Milner's virial values into Ghosh's formula.

Part I.

THE problem of the anomaly of strong electrolytes has for the last thirty years or more, been one of the most baffling in physical chemistry. Some years ago J. C. Ghosh † put forward a theory which at the time appeared to go a long way towards the complete solution of this problem. The theory is based on the fundamental assumption that the solutions of strong electrolytes are completely ionized at all dilutions and that the ions attract each other according to the inverse square law as point charges placed in a continuous medium having the dielectric constant of the solvent. A very large amount of experimental evidence is adduced in support of the formulæ put forward. To arrive at these, Ghosh introduces a number of subsidiary hypotheses which would have to be proved to make the theory complete. The same formulæ are claimed to govern the variations of molecular conductivity with dilution in the case of all electrolytes of similar type, and these formulæ contain only

* Communicated by the Author.

† Chem. Soc. Journ. cxiii. p. 449; *ibid.* p. 627; *ibid.* p. 707; *ibid.* p. 790 (1918).

constants that are independent of the conductivity measurements themselves. The validity of this claim was disputed by J. R. Partington *, who also criticised the theory in other respects. A much more serious objection was raised by D. L. Chapman and H. J. George †, who pointed to the large number of subsidiary hypotheses and showed that one of Ghosh's principal formulæ was in contradiction with the hypothesis stated. This fact had been noticed by me about nine months previous to the publication of Chapman and George's note. On referring the matter to Dr. Ghosh, I was informed that the formulæ were modelled on the well-known equation of state of Dieterici ‡. I have since that time been engaged in a careful study of Ghosh's theory in the hope that it might be possible to vindicate it completely by a slight adjustment of the subsidiary hypotheses and perhaps to establish these more firmly. As will be seen the hypotheses may be adjusted to fit the formulæ, and on the whole the hypotheses are strengthened thereby. The attempts, however, to establish these have not been successful, and the conclusion has been reached that it is necessary to abandon the hypothesis of a uniform dielectric medium. Since the present study was undertaken a considerable number of papers has been published dealing with Ghosh's theory §. As, however, these treat the matter from different points of view from those adopted in this paper, it will be sufficient to refer to them, where their conclusions touch those reached in the present instance.

Degree of association of a completely ionized electrolyte in a uniform dielectric medium.

The fundamental assumption of the complete dissociation of strong electrolytes is not new. It has formed the basis of theoretical work by W. Sutherland ‖, by S. R. Milner ¶, by P. Hertz **, by N. Bjerrum ††. A clear distinction between

* Trans. Faraday Soc. xv. pp. 110 & 162 (1919).

† Phil. Mag. xli. p. 799 (1921).

‡ *Ann. Phys.* [4] v. p. 51 (1901).

§ J. Kendall, Am. Chem. Soc. Journ. xliv. pp. 717–738 (1922); Sv. Arrhenius, *Zeit. Phys. Chem.* c. pp. 9–25 (1922); C. A. Kraus, Amer. Chem. Soc. Journ. xliii. p. 2514 (1921); H. Kallmann, *Zeit. Phys. Chem.* xcviii. p. 433 (1921); and others.

‖ Phil. Mag. iii. p. 161 (1902); *ibid.* xii. p. 1 (1906); *ibid.* xiv. p. 1. (1908).

¶ Phil. Mag. xxiii. p. 551 (1912); *ibid.* xxv. p. 742 (1913); *ibid* xxxv. pp. 214 & 352 (1918).

** *Ann. d. Physik,* xxxvii. p. 1 (1912).

†† *Zeit. Elektrochem.* xxiv. p. 321 (1918).

complete ionization and complete dissociation should be made. As the result of modern work on crystal structure the view has become very generally accepted that crystals of salts such as sodium chloride consist of ions which are held together by electrical forces. When such crystals are dissolved in water no reason is apparent why any other mode of combination than holds in the solid state should come into play, *i. e.* the substance would remain completely ionized. The question would, however, still be open whether any considerable fraction of oppositely charged ions would be held by their charges in such close proximity that they would form doublets which would be indistinguishable from molecules. If this were so, as we have every reason to believe is the case with salt molecules in the condition of vapour, most of the properties of the substance, such as its gaseous pressure and the absence of electrical conductivity under the action of moderate potential gradients, will be the same as those of an undissociated substance. We should therefore term it "undissociated," and yet it would be completely ionized inasmuch as the mutual attraction of the ions would obey comparatively simple electrical laws and its properties would follow from general considerations based on such laws.

The question can be quantitatively approached by means of a theorem of the dynamical theory of gases due to Boltzmann, which Milner has also made the basis of his work on the anomaly of strong electrolytes in the papers referred to. We employ the theorem in the form in which it is stated by Milner, *loc. cit.* 1912, p. 553, viz. if p_W and p_W' stand for the probabilities of the occurrence of configurations of which the potential energies are W_1 and W_2 respectively in a system of molecules where intermolecular forces exist and p, p' stand for the probabilities of the occurrence of the same configurations in the same system of molecules if there were no intermolecular forces, then :

$$\frac{p_W}{p_W'} = \frac{p}{p'} e^{-\frac{W_1 - W_2}{\omega}}, \quad \ldots \quad (1)$$

where ω stands for the most probable value of the kinetic energy of a molecule. If R is the gas constant, T the absolute temperature, and N the number of molecules in a mol (Avogadro's constant), then

$$\omega = \frac{RT}{N} \ldots \quad (2)$$

This theorem may be made the base of an approximate

calculation to determine to what extent the probability that oppositely charged ions should be found in close proximity to each other, is affected by their charges, if we confine our attention to one pair of ions. We imagine the volume available to be equal to the average volume occupied by a pair of ions in the solution and assume the space to be of spherical shape with one of the ions fixed in the centre. Further, we assume that in the absence of any attractive force between the two ions all positions in small volumes of equal size in the sphere are equally probable for the second ion.

Before proceeding further it will, however, be necessary to consider more carefully the law according to which ions in a solution attract each other. We shall provisionally assume a uniform dielectric constant of the medium as was done by Ghosh, returning to the discussion of this assumption later on. As regards a comprehensive law according to which ions attract each other, we have particularly the researches of M. Born and A. Landé on the compressibility of salt crystals to refer to *. On these Born bases a theory of ionic structure leading to an expression for the potential energy W of a pair of attracting ions, of the form :

$$W = -\frac{a}{r} - \frac{b}{r^5} + \frac{c}{r^9}. \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Herein r is the distance between the centres of the two ions ; the constant a has the value which would make the first term express the potential energy of the pair of ions if they were point charges, *i. e.*

$$a = \frac{E^2}{\epsilon}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where E is the elementary ionic charge and ϵ the dielectric constant of the medium. The second term is small, and may be neglected so that we may make

$$b = 0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

and the third may be expressed in terms of the distance which the centres of the ions would have if they were in equilibrium. This distance, which equals approximately the sum of the ionic radii in the crystalline state, we indicate

* *Verh. d. D. Phys. Ges.* xix. p. 210 (1918) ; *ibid.* xx. p. 232 (1918).

by d_0. Thus we have

$$\frac{dW}{dr} = \frac{a}{d_0^2} - 9\frac{c}{d_0^{10}} = 0 \quad (r = d_0),$$

or

$$c = \frac{ad_0^8}{9}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

Substituting (4), (5), and (6) into (3) we obtain

$$W = -\frac{E^2}{\epsilon r} + \frac{E^2 d_0^8}{9\epsilon r^9}. \quad . \quad . \quad . \quad . \quad . \quad (3a)$$

The diameters of various ions have values for which we may refer to a table given by W. L. Bragg *. For our present purpose it will be sufficient to take as a representative value say $d = 3 \times 10^{-8}$ cm.

We now come to the consideration of a concrete case. We imagine the solution to be of tenth normal strength. In this case the average volume occupied by a pair of ions will be $\frac{10000}{N}$ c.c., N being Avogadro's number, and this would be the volume of the sphere referred to above. The radius of such a sphere is $15{\cdot}85 \times 10^{-8}$ cm. If no forces existed between the ions, but the other conditions we have postulated held, we calculate for the ratio of the two probabilities (1) that the ions are in contact, that is that their centres have a distance $r_1 = 3 \times 10^{-8}$ cm., and (2) that the second ion is somewhere near the circumference of the sphere, say that the distance of its centre from the centre of the sphere is $r_2 = 14 \times 10^{-8}$, the value :

$$\frac{p}{p'} = \frac{3^2}{14^2} = 0{\cdot}046. \quad . \quad . \quad . \quad . \quad . \quad (7)$$

If, however, the ions attract each other according to Born and Landé's law (3 *a*), we find their potential energies, making $d_0 = 3 \times 10^{-8}$ cm., $E = 4{\cdot}7 \times 10^{-10}$ e.s.u., $\epsilon = 80$, to be according to equation (3 *a*) respectively

$$W_1 = -8{\cdot}1818 \times 10^{-14} \text{ erg}; \quad . \quad . \quad . \quad . \quad (8)$$

$$W_2 = -1{\cdot}9723 \times 10^{-14} \text{ erg}. \quad . \quad . \quad . \quad . \quad (9)$$

Making $R = 8{\cdot}317 \times 10^7$ ergs per degree and mol, $T = 17°$ C. $= 290°$ abs., $N = 6{\cdot}16 \times 10^{23}$, and substituting the values given by (2), (7), (8), and (9) into Boltzmann's equation (1),

* Phil. Mag. xl. p. 179 (1920).

we now calculate for the relative probabilities of the two configurations considered the ratio :

$$p_W : p_W' = 0{\cdot}225,$$

i. e., the probability that the centres of the ions are in contact, *i. e.* at a distance of 3×10^{-8} cm. is in spite of the attractive forces only a little more than one-fifth of the probability that they are at the limit assumed, *i. e.* at a distance of about 14×10^{-8} cm. from each other. In this case the probability that the second ion is present in a given small volume in contact with the first is only 4·9 times as great as that it is in an element of volume of the same size near the circumference of the sphere considered. By making $W_2 = 0$ we can calculate that in an unlimited space the probability of the second ion being in contact with the first in a given small volume is 8·1 times as great as that it is in a small volume of equal size at such places where the attraction is negligible. On the other hand, if we make $\epsilon = 1$, and give to T so high a value even as T 1273° corresponding to the case of NaCl vapour at 1000° C., we find for the above ratio the value $3{\cdot}4 \times 10^{16}$. From this we conclude that even for very large volumes, *i. e.* at low pressures, the probability of the ions being in close proximity is overwhelmingly much greater than that they would be free from each other, and we should expect NaCl vapour to behave as an undissociated substance even although we assume it to be completely ionized. The absence of dissociation in the sense generally understood can therefore in this case not be employed as an argument against complete ionization as has been done in a recent discussion (R. M. Caven, Nature, cviii. p. 210 (1921)).

Approximate range of simple inverse square law for attraction of ions.

Another matter which is of interest is to know to what extent a law such as that of Born and Landé invalidates the use of the inverse square law. A simple calculation yields the result that if the ions are distributed uniformly in a solution of $\frac{N}{10}$ strength the volume of that part of it, in which the potential calculated according to Born and Landé differs by more than one per cent. from the value given by the inverse square law, is only 1·67 per cent. of the total volume of the solution. If we wish to find the probability that the attraction in the case of a given pair of ions does

not show a greater deviation from the inverse square law than that mentioned above, we should have to allow for the fact that positions of proximity are more probable than others according to Boltzmann's law. Here again as above, we can only arrive at an approximation by considering one of the ions fixed at the centre of a sphere the volume of which is equal to the average volume available for the pair of ions in the solution and the other ion free to move in this sphere. Even in this comparatively simple case the calculation is long and troublesome and leads to the result that the probability that the potential of the pair of ions would differ by more than one per cent. from that following from the inverse square law is about 3·7 per cent.

We may summarize the preceding considerations by saying: On the assumption that the solvent separating ions has a constant specific inductive capacity equal to that of the liquid in bulk, the hypothesis of complete ionization in the case of salts such as sodium chloride is well established, electrical association say in N/10 solution would be small, and the preponderating proportion of ions would be subject to the inverse square law of electrical action.

Ghosh's virial formula. Heat of dilution of an electrolyte conforming to Milner's or Ghosh's osmotic pressure formula.

As already stated, Ghosh requires for the purpose of his theory to evaluate the work which must be done to remove the ions arising from say a mol of salt in solution, from the sphere of each other's action. Milner (*loc. cit.* 1912) has calculated this quantity on the basis of Boltzmann's theorem for a binary electrolyte subject to the assumptions we have made above. In the case of forces obeying the inverse square law, this function is the same as the negative value of the one introduced into theoretical physics by Clausius under the name of the virial, and Milner terms it the attractive virial of the solution. The calculation is an extremely long and difficult one. Milner has applied the result to calculate the osmotic pressure, and hence the lowering of freezing points of binary electrolytes by making use of the Clausius' Virial theorem, according to which pressure times volume equals two-thirds kinetic energy minus one-third attractive virial. Milner discusses the application of this theorem to the osmotic pressure of a solution and supplies a new proof based on thermodynamical reasoning (*loc. cit.* xxv. p. 747). This proof however, like the dynamical one of

Clausius, applies only to gases for it contains two assumptions which do not hold for liquids. First, the temperature-coefficient of the dielectric constant is assumed to be zero, and Milner himself draws attention to the fact that this does not hold for liquids. Secondly, it contains the assumption that the difference in internal energy brought about in a solution by dilution is the result solely of overcoming the electrical forces between the ions, and that therefore the cooling of the solution should be the exact equivalent of this work. It is, however, doubtful whether the process of dilution in the case of an electrolytic solution does not essentially combine with the cooling effect just referred to, a production of heat of a similar order of magnitude. The matter is most readily understood by proceeding in the opposite direction to that adopted by Milner.

For the virial W of a solution Milner deduces a result of the general form :

$$W = RTh\,f(h), \quad . \quad . \quad . \quad . \quad . \quad (10)$$

where f is a function of h for which values may be obtained from a table supplied (p. 749, *loc. cit.*) and h is a simple function of the product $V^{1\,3}T\epsilon$, V being the dilution (*i. e.* volume per mol solute), T the absolute temperature, and ϵ the dielectric constant. The result may therefore be still more simply expressed as

$$W = T\mathfrak{F}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

where $\mathfrak{F}$ is a function of the product

$$\sigma = \epsilon V^{1/3}T, \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

ϵ being considered a function of T only. In conjunction with the Clausius' virial theorem, Milner's result (10) applied to two mols of ions thus yields the relation :

$$PV = 2RT - \tfrac{1}{3}RTh\,f(h), \quad . \quad . \quad . \quad . \quad (13)$$

or more simply according to (11),

$$PV = 2RT - \tfrac{1}{3}T\mathfrak{F}. \quad . \quad . \quad . \quad (13\,a)$$

As explained above Milner further assumes that

$$\frac{\partial U}{\partial V} = -\frac{\partial W}{\partial V}, \quad . \quad . \quad . \quad . \quad . \quad (14)$$

where U is the internal energy of the solution. The two

equations (13 *a*) and (14) combined are, however, compatible with the thermodynamical relation also utilized by Milner :

$$\frac{\partial U}{\partial V} = T\frac{\partial P}{\partial T} - P^* \quad . \quad . \quad . \quad . \quad . \quad (15)$$

only when $\frac{d\epsilon}{dT} = 0$, for, substituting the values for P and $\frac{\partial P}{\partial T}$, which follow from (13 *a*), into (15), and comparing the result with that following from equations (11) and (12) for $\frac{\partial W}{\partial V}$, we find after some transformations :

$$\frac{\partial U}{\partial V} = -\frac{\partial W}{\partial V} - \frac{T^3}{3V^{2/3}}\frac{d\mathfrak{F}}{d\sigma}\frac{d\epsilon}{dT},$$

and ultimately :

$$\frac{\partial U}{\partial V} = -\frac{\partial W}{\partial V}\left(1 + \frac{T}{\epsilon}\frac{d\epsilon}{dT}\right), \quad . \quad . \quad . \quad . \quad (16)$$

in place of the relation (14), assumed by Milner, which it would thus become necessary to drop in any attempt to apply his result to solutions. By integrating equation (16) we obtain for the heat of dilution H to zero concentration, the result :

$$H = U - U_\infty = -W\left(1 + \frac{T}{\epsilon}\frac{d\epsilon}{dT}\right). \quad . \quad . \quad . \quad (17)$$

A test of this formula by available heat data does not lead to satisfactory agreement, and we conclude therefore that an expression for the osmotic pressure of a solution of the general form given by equation (13), in conjunction with (12), which, as will be seen shortly, is common both to Milner and Ghosh, can have only approximate validity.

* In a recent paper (Phil. Mag. xliii. p. 625 (1922)) B. Cavanagh criticises Milner's application of the virial theorem to solutions. He further remarks, evidently with reference to Milner's use of the above thermodynamical equation (15), "The demonstration is open to grave doubts, for it depends on treating the solute thermodynamically as an independent system whose external pressure is the osmotic pressure of the solution." This latter criticsm can however not be admitted, for equation (15) is a direct consequence of the quite general Gibbs-Helmholtz equation and the very general relation $\frac{\partial A}{\partial V} = -P$, where A is the free energy, P the osmotic pressure, and V the dilution of a solution containing a given amount of solute and separated from pure solvent by a semipermeable membrane.

Ghosh's Virial Expression.—Like Milner, Ghosh applies the Clausius virial theorem, but in place of the virial deduced on the basis of the kinetic theory makes use of an energy term of an altogether different kind. He assumes that a mean disposition of the oppositely charged ions exists in solution and, further, that this disposition conforms to some simple pattern in space, analogous to the arrangement of the atoms in a solid crystal. Thus, in the case of a univalent binary salt, the mean positions of the ions are supposed to form a cubic space lattice, the distance r between two adjacent ions being calculated to be $r = \sqrt[3]{\frac{V}{2N}}$, where V is the dilution and N Avogadro's number. In the case of salts such as barium chloride a disposition such as that of the atoms of a fluorspar crystal is assumed; the distance between a barium and a chlorine ion is calculated to be $\frac{\sqrt{3}}{2} r$, where r has the same value as in the case of binary salts. With salts such as magnesium sulphate adjacent ions are supposed to approach each other more closely than in the case of univalent salts, the distances in the latter and former cases standing in the ratio $\sqrt[3]{4}$. In every case Ghosh calculates the work that is necessary to separate adjacent ions constituting an original molecule from each other and multiplies the result by the number of original molecules in the solution. If we indicate this work per gram-molecule of solute by G, Ghosh obtains in the case of a univalent binary electrolyte

$$G = \frac{NE^2 \sqrt[3]{2N}}{\epsilon \sqrt[3]{V}} = \frac{\sqrt[3]{2}}{10} \frac{F^2}{\epsilon N^{2/3}} \sqrt[3]{C}, \quad . \quad . \quad (18)$$

where N is Avogadro's constant, E the electronic charge, ϵ the dielectric constant, F Faraday's constant $9654 \times 3 \times 10^{10}$ e.s.u. per g-equiv., and C the concentration measured in mols per litre. For an electrolyte such as barium chloride the result is put forward:

$$G = \frac{3NE^2 2 \sqrt[3]{2N}}{\epsilon \sqrt{3} \sqrt[3]{V}} \cdot \cdot \cdot \cdot \cdot \cdot \quad (19)$$

This result would, however, correspond to an arrangement in which the distance of the chlorine atoms from each other is the same as their distance from the barium atom and is about four per cent. smaller than the value following from the arrangement postulated by Ghosh. Ghosh also erroneously defines the quantity which we have indicated by G as the work which is necessary to separate completely the

ions constituting an original gram-molecule of solute, *i. e.* $n \times N$ ions, n being the number of ions into which a molecule dissociates. This definition does not, however, agree with the arrangements of ions assumed, G being considerably smaller in every case. Thus for the lattice postulated for a univalent binary electrolyte, the energy necessary to separate the ions has been calculated by Madelung * in connexion with other researches, and is nearly 75 per cent. greater than G. The discrepancy between the values of G assumed by Ghosh and the interpretation given to them has also been pointed out by Kraus †, and a value agreeing with that previously given by Madelung has been quite recently calculated by Kendall ‡ for univalent binary electrolytes.

In further dealing with Ghosh's theory it is thus evident that we must drop altogether the idea of an actual arrangement of the ions in the solution like that of the assumed space lattices. Indeed, it is plain that if a literal interpretation were given to these space lattices, the electrolytic solution or parts of it, should have axes, that is definite directions in space in which its properties were different from those in other directions. Two alternatives appear to be open to us. We may either consider the quantities G to be purely empirical energy expressions which when substituted into the Clausius virial equation yield an expression for the osmotic pressure agreeing, as will be shown below, within certain ranges with deductions from experiment; or else we may adhere to the interpretation of G as the work necessary to separate completely one gram-molecule of ions. This could be done, for example, by assuming that groups of ions constituting an original molecule were arranged in the manner required by Ghosh's space lattices, but that these groups had a random distribution among themselves. The artificiality of such explanations may justly be urged as an objection to them, and before proceeding further along the lines opened up by such theories, the rôle of the dielectric to which reference will be made below will no doubt have to be more closely scrutinized. In the meantime we shall probably do best to consider Ghosh's space lattices merely as theoretical "distributions of reference" the axes of which may be placed in arbitrary directions and which serve to yield units of length that are empirically important in the presentation of the free energy and conductivity relations of an electrolytic solution.

* *Phys. Zeit.* xix. p. 524 (1918).
† *Loc. cit.* ‡ *Loc. cit.*

Comparison of the deductions from Ghosh's and Milner's osmotic pressure formulæ, with experimental results.

Ghosh's expression for osmotic pressure has the same general form as Milner's, which is represented by equations (13 *a*) and (12), and the disagreement between equation (17) and experiment referred to above thus shows that both can at best only claim limited validity. This validity we now come to examine. A comparison between theory and experiment is best made by utilizing van't Hoff's factor i which is defined by

$$PV = iRT.$$

By the aid of this equation we find on making the virial in the Clausius theorem applied to an original mol of electrolyte respectively equal to $RThf(h)$ according to Milner, or equal to G according to Ghosh, that according to Milner :

$$2-i = \tfrac{1}{3}hf(h); \quad . \quad . \quad . \quad . \quad . \quad (20)$$

whereas according to Ghosh :

$$2-i = \tfrac{1}{3}\frac{G}{RT}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

This equation (21) yields for univalent binary electrolytes (on substituting the value (18) for G and making $\epsilon = 87{\cdot}5$ at 0°, $T = 273°$, $R = 8{\cdot}32 \times 10^7$ ergs per degree per mol)

$$2-i = 0{\cdot}2446\sqrt[3]{C},$$

C being the concentration in mols per litre.

The following Table I. shows in the second and third lines the values of $2-i$ at various concentrations calculated according to Milner and to Ghosh respectively, and in the fourth line the ratio between the two. It will be seen that at about N/100 concentration the two are practically identical, whereas at lower concentrations Milner's values are smaller, and at higher concentrations greater than those of Ghosh. Milner's table (p. 749, *loc. cit.*), giving values for $\frac{1}{3}hf(h) = 2-i$ at varying concentrations, may also be applied to binary bivalent electrolytes such as $CuSO_4$, for a study of Milner's expression for h shows that binary bivalent electrolytes should, according to the theory, have the same value for $2-i$ as univalent ones, if their concentration is one sixty-fourth of the concentration of the univalent electrolyte. For bivalent electrolytes Milner's virial is the same as Ghosh's function at a concentration of about 3·1 thousandth normal and is smaller at lower concentrations.

Table I.

Concentrations (mols per litre)	0·002	0·005	0·01	0·02	0·05	0·1	0·2	0·5
$2-i=\frac{1}{3}hf(h)$, acc. to Milner.....	0·0245	0·038	0·053	0·074	0·115	0·162	0·224	...
$2-i=\frac{1}{3}G/RT$, acc. to Ghosh	0·031	0·042	0·0527	0·0664	0·0901	0·1135	0·1431	0·1942
Ratio $(2-i)$ Milner to $(2-i)$ Ghosh	0·78	0·89	0·99	1·094	1·255	1·40	1·51	...
Percentage error in $2-i$ for 1/1000° ...	...	...	93	42	11	4	2	$\frac{1}{2}$
an observational error of 1/10,000°...	84	29	9	4	1	0·4	...	...
$(2-i)$ Experimental mean acc. to Ghosh	...	...	0·058 (9%)	0·072 (4%)	0·100 (12%)	0·120 (17·5%)	0·151 (14%)	0·195 (12%)
$(2-i)$ Experimental mean acc. to Noyes and Falk	...	...	0·063	0·075	0·113	0·135	0·163	0·196
$(2-i)$ for KCl (Adams)	0·032	0·039	0·057	0·078	0·112	0·159	...	...
$(2-i)$ for KNO_3 (Adams)	0·033	0·042	0·063	0·092	0·152	0·213	...	...

Table II.

Values of $\frac{2-i}{\sqrt[3]{C}}$ for KIO_3 calculated from experiments by Hall and Harkins.

Theoretical value according to Ghosh equal 0·2446.

Concentrations (mols per litre)	0·00274	0·00506	0·00903	0·01358	0·03244	0·09380
$\frac{2-i}{\sqrt[3]{C}}$	0·396	0·322	0·386	0·393	0·419	0·506

As regards the experimental determination of $2-i$ we are restricted almost entirely to the results of cryoscopic measurements. Even here the results must be considered doubtful at low concentrations. Recent American workers have claimed an accuracy of 1/10000 degree for their measurements. We give accordingly in the fifth and sixth lines of Table I. the percentage variations in $2-i$ that correspond to an observational error in the cryoscopic measurements of 1/1000 and 1/10000 degree respectively. It will be realized that the test as here applied is a severe one, and also that there is little purpose in applying it to concentrations much below one-hundredth normal. Starting from the assumption that the values of i are the same for all uni-univalent electrolytes, Ghosh has calculated the mean of the principal available data for halogen salts, and his results are given in line 7 of Table I. Underneath each figure is given the maximum percentage variation from this mean in the case of the mean of any one given salt. In line 8 are given similar results that had been previously tabulated by Noyes and Falk*. In line 9 are given some more recent values for $2-i$ for KCl determined by L. H. Adams, and in line 10 similar values for KNO_3†. In all calculations the molecular lowering for water has been taken as 1·858 (degree, litre, mol).

In Table II. some results are given for $2-i/\sqrt[3]{C}$ which I have calculated from experiments by R. E. Hall and W. D. Harkins‡ on KIO_3 and which may be compared with the value 0·2446 following from Ghosh's formula. The conclusion to be drawn from these results is that there are considerable differences between the i values of various univalent binary electrolytes, but that in the case of the halogen salts, the differences are comparatively small and some of the values agree well with the formulæ of Milner and of Ghosh.

When we come to binary bivalent electrolytes the formula of Ghosh gives the result $(2-i):\sqrt[3]{C}=1{\cdot}553$, and as already explained Milner's formula yields practically the same value for a concentration of about three thousandth normal. The table supplied by Milner does not give values for $2-i$ at higher concentrations, whereas at lower concentrations the above expression would be smaller than 1·553.

In Table III. are given some values for $(2-i)/\sqrt[3]{C}$ for various salts, most of which have been calculated from data

* J. Amer. Chem. Soc. xxxii. p. 1027 (1910).
† *Ibid.* xxxvii. p. 481 (1915).
‡ *Ibid.* xxxviii, p. 2658 (1916).

Table III.

Values of $\frac{2-i}{\sqrt[3]{C}}$ calculated from experimental data. Theoretical value according to Ghosh equal 1·553.

$MgSO_4$	conc.	0·00320	0·01310	0·02668	0·04878	0·09566	0·2210	...
	$\frac{2-i}{\sqrt[3]{C}}$	2·16	2·18	2·15	1·90	1·70	1·46	...
$CdSO_4$	conc.	0·000704	0·001511	0·002685	...	...	...	...
	$\frac{2-i}{\sqrt[3]{C}}$	2·20	2·47	2·58	...	...	...	...
$CuSO_4$	conc.	0·001521	0·003441	0·01463	0·1051	0·2074	0·4034	0·8894
	$\frac{2-i}{\sqrt[3]{C}}$	2·63	2·77	2·47	1·64	1·60	1·37	1·09
$ZnSO_4$	conc.	0·001695	0·002850	0·01391	0·1278	0·3112	1·001	...
	$\frac{2-i}{\sqrt[3]{C}}$	2·23	2·52	2·37	1·59	1·35	0·99	...

Table IV.

Values of $\frac{3-i}{\sqrt[3]{C}}$ calculated from various experiments. Theoretical value according to Ghosh equal 0·8475.

$BaCl_2$ acc. to Hall and Harkins	conc.	0·00273	0·00534	0·01142	0·02982	0·05611	0·81136	...
	$\frac{3-i}{\sqrt[3]{C}}$	1·153	1·248	1·246	1·188	1·116	0·978	...
$CoCl_2$ acc. to Hall and Harkins	conc.	0·00209	0·01008	0·02251	0·05475	0·1256	0·2772	0·4217
	$\frac{3-i}{\sqrt[3]{C}}$	1·107	1·161	1·090	0·942	0·742	0·452	0·273
Experimental Mean adopted by Ghosh for electrolytes of type	conc.	...	...	0·02	0·05	0·10	..	...
	$\frac{3-i}{\sqrt[3]{C}}$			0·843	0·830	0·810		

collected by G. N. Lewis and G. A. Linhart *. The experiments are mainly due to Hall and Harkins † and to Hausrath ‡.

For ternary electrolytes such as $BaCl_2$ Ghosh's formula leads to the result $(3-i)/\sqrt[3]{C}=0{\cdot}8475$. Values for this quantity are assembled in Table IV. which have been calculated from experiments by Hall and Harkins and also from numbers given by Ghosh as representative mean values for various electrolytes.

Most of the results in Tables III. and IV. can only be said to agree with Ghosh's formulæ as regards order of magnitude, though the agreement between the mean of experimental values calculated by Ghosh for ternary electrolytes like $BaCl_2$ and his formula is close.

For electrolytes such as K_2SO_4 the values for $(3-i)/\sqrt[3]{C}$ differ considerably from the number 0·8475. They vary from about 1·3 to 1·7.

In conclusion, the application of Ghosh's formula to ebullioscopic results must be referred to. Confining ourselves to aqueous solutions of univalent binary electrolytes and therefore making $T=373°$, $\epsilon=52{\cdot}6$, Ghosh's formula leads to $(2-i)/\sqrt[3]{C}=0{\cdot}306$; whereas a selection of results calculated for the same quantity from various experiments was found to vary from about 0·18 to about 0·58.

We may summarize the conclusions to be drawn from the results given by saying that the figures do not support the contention that van't Hoff's factor is solely dependent on the influences we have considered, and it cannot be maintained that it has the same value for all equimolecular solutions of electrolytes of similar type. In the cases reviewed the values following for $(n-i)$ from Ghosh's formulæ agree in order of magnitude with the available experimental data. As regards mean values for univalent binary chlorides in aqueous solution, the agreement between $(2-i)$ calculated from cryoscopic measurements and the same quantity derived from Ghosh's formula is good. Particularly at higher concentrations the agreement is better than that attained by Milner's formula, and empirically Ghosh's formula will therefore probably be preferred to the more elaborate theoretical one of Milner.

* J. Amer. Chem. Soc. xli. p. 1951 (1919).
† *Ibid.* xxxviii. p. 2672 (1916).
‡ *Ann. Phys.* ix. p. 522 (1902).

X. *On the Problem of two Moving Charges in its connexion with the Atomic Theory.* *By* S. Boguslawski, *Professor at Moscow University* *.

I. *Motion of two Charges with Masses varying according to the Lorentz-Einstein Law.*

1. THE investigations of A. Sommerfeld have shown that the variation of mass of an electron with its velocity explains the structure of spectral lines of hydrogen and helium. Thus the facts observed in spectroscopy have come to support the accuracy of the Lorentz-Einstein formula for the variation of mass with velocity.

On the other hand, C. G. Darwin † and A. Sommerfeld ‡ carried out a closer investigation of the orbits of electrons moving according to this law under the attraction of a fixed positive charge, which shows that in certain circumstances the electron ought to describe spiral trajectories round the positive nucleus. There is no experimental evidence in favour of these orbits, and there is a considerable amount of indirect evidence which leads one to believe that these orbits do not occur in reality. A purely formal explanation of the unreality of those orbits is readily supplied by the quantum theory, the angular momentum for them being so small as not to be admitted by Bohr's theory of atoms. But it is doubtful how far this explanation may be considered as sufficient. It may be claimed that the quantum theory selects only those orbits which are in a certain sense stable, and does not deny altogether the possibility of other unstable orbits. From the latter point of view, the investigations of C. G. Darwin and A. Sommerfeld dealing with spiral orbits have not lost their interest.

There is a point in these investigations which makes them appear incomplete, and their result not so convincing as is desirable: I mean the fact that both these authors consider the central charge as fixed in space. This supposition gives a result of a high degree of accuracy so long as the mass of the central charge is large compared with the mass of the charge revolving round it. But this last condition ceases to be fulfilled from a certain moment while the charge

* Communicated by the Author.

† C. G. Darwin, Phil. Mag. (6) vol. xxv. p. 201 (1913).

‡ A. Sommerfeld, *Ann. d. Phys.* 1916; also 'Atombau und Spektrallinien,' 2nd edition, p. 324 (1921).

is approaching the central kernel along a spiral trajectory. As the distance between them decreases, the velocity and the mass of the moving charge must both increase so that the latter may become comparable with that of the central charge. It is true that this will only occur when the distance between the charges is comparable with the diameter of the electron. In this case our knowledge of the dynamics of any masses becomes uncertain. As far as experience on scattering of α-rays goes, Coulomb's law seems to hold good even for distances of this order of magnitude. At lower distances, however, this law probably ceases to be true, as is shown by later experiments of Sir E. Rutherford on the production of H-rays by collision of α-particles with atoms. For this reason it is doubtful whether theoretical orbits which, like the spiral ones, approach very closely the central nucleus give an accurate description of real facts. But all the same, even if the theory does not here find immediate application, the theoretical investigation of an idealized problem does not lose its interest, and may help in some way to find the correct solution. From this standpoint, I thought it useful to study the motion of two charges whose masses vary according to the law given by the theory of relativity. I could not find out whether anybody has already studied this problem. But owing to the state of things in our country, we were so badly supplied with literature that possibly I am ill-informed.

The result of this study will be to confirm the results previously obtained by C. G. Darwin and A. Sommerfeld, in the sense that by taking into account the motion of the central nucleus, we still obtain spiral trajectories in relative motion of the two charges when the relative angular momentum is sufficiently small.

2. To find the solution of our problem we shall apply the method of the differential equation of Hamilton-Jacobi. This method seems to be the most appropriate for purposes of atom dynamics, because it gives the result in a form to which the principles of the quantum theory can be immediately applied. But also in cases where no use of quantum theory is made, the method can be recommended for its compactness.

Let e_1 and e_2 denote two charges, whose masses for infinitely small velocities are respectively m_1 and m_2. Let x_1, y_1, z_1 be the coordinates of the first charge; x_2, y_2, z_2

those of the second. We put, according to the general use,

$$\beta_1^2 = \frac{1}{c^2}(\dot{x}_1^2 + \dot{y}_1^2 + \dot{z}_1^2), \quad \beta_2^2 = \frac{1}{c^2}(\dot{x}_2^2 + \dot{y}_2^2 + \dot{z}_2^2),$$

where c is the velocity of light.

The Lagrangean function of our problem is

$$L = -m_1c^2\sqrt{1-\beta_1^2} - m_2c^2\sqrt{1-\beta_2^2} - \frac{e_1e_2}{r} \quad . \quad . \quad (1)$$

with

$$r^2 = (x_1-x_2)^2 + (y_1-y_2)^2 + (z_1-z_2)^2.$$

The impulses of the two masses will be

$$p_{x_1} = \frac{\partial L}{\partial \dot{x}_1} = \frac{m_1\dot{x}_1}{\sqrt{1-\beta_1^2}}, \ldots, \quad p_{x_2} = \frac{\partial L}{\partial \dot{x}_2} = \frac{m_2\dot{x}_2}{\sqrt{1-\beta_2^2}}, \ldots . \quad (2)$$

The Hamiltonian function of the problem which expresses the total energy of the system is

$$H = \Sigma\dot{x}_1 p_{x_1} + \Sigma\dot{x}_2 p_{x_2} - L, \quad . \quad . \quad . \quad . \quad (3)$$

where the symbol Σ means a summation over the three coordinates. Making use of (1) and (2), and expressing the velocities in terms of the impulses, we bring (3) in the form

$$H = c\sqrt{m_1^2c^2 + \Sigma p_{x_1}^2} + c\sqrt{m_2^2c^2 + \Sigma p_{x_2}^2} + \frac{e_1e_2}{r}. \quad . \quad (4)$$

If, now, we suppose $p_{x_1}, \ldots, p_{z_2}$ to be the six partial differential coefficients of a function W corresponding to the six independent variables $x_1, \ldots z_2$, and put the right side of (4) equal to a constant α_1', we shall obtain the partial differential equation of Hamilton-Jacobi, on whose solution depends the solution of our problem. The constant α_1' is the total energy of our system, consisting of the kinetic energy of both masses, their mutual potential energy as expressed by the last term in (4), and of the internal energy stored in both masses, which according to the theory of relativity amounts to m_1c^2 and m_2c^2 respectively for the two masses.

Transforming the differential equation in order to get rid of the irrationalities, we can write it in the form

$$c^4(m_1^2-m_2^2)^2 + \Sigma p_{x_1}^2 - \Sigma p_{x_2}^2)^2 + 2c^2(m_1^2-m_2^2)(\Sigma p_{x_1}^2 - \Sigma p_{x_2}^2)$$
$$+ \frac{1}{c^2}\left(\alpha_1' - \frac{e_1e_2}{r}\right)^2 \left\{ \frac{1}{c^2}\left(\alpha_1' - \frac{e_1e_2}{r}\right)^2 \right.$$
$$\left. - 2(m_1^2+m_2^2)c^2 - 2(\Sigma p_{x_1}^2 + \Sigma p_{x_2}^2) \right\} = 0. \quad . \quad (5)$$

This is the equation that we have now to solve.

3. We introduce new variables by putting

$$\left.\begin{aligned} x &= x_1 - x_2, & \xi &= x_1 + x_2, \\ y &= y_1 - y_2, & \eta &= y_1 + y_2, \\ z &= z_1 - z_2, & \zeta &= z_1 + z_2, \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (6)$$

where x, y, z are obviously the relative coordinates of the first mass as referred to the second. The partial differential coefficients $p_{x_1}, \ldots p_{z_2}$ will then be expressed in terms of the new ones $p_x, \ldots p_\zeta$ as follows :—

$$\left.\begin{aligned} p_{x_1} &= p_\xi + p_x, & p_{x_2} &= p_\xi - p_x, \\ p_{y_1} &= p_\eta + p_y, & p_{y_2} &= p_\eta - p_y, \\ p_{z_1} &= p_\zeta + p_z, & p_{z_2} &= p_\zeta - p_z. \end{aligned}\right\} \quad . \quad . \quad . \quad (7)$$

Solving these equations, we shall have

$$p_\xi = \tfrac{1}{2}(p_{x_1} + p_{x_2}, \qquad p_x = \tfrac{1}{2}(p_{x_1} - p_{x_2}) ; \quad . \quad . \quad . \quad (8)$$

and similarly for the two other axes.

It follows from (7)

$$\left.\begin{aligned} \Sigma p_{x_1}^2 + \Sigma p_{x_2}^2 &= 2(\Sigma p_\xi^2 + \Sigma p_x^2), \\ \Sigma p_{x_1}^2 - \Sigma p_{x_2}^2 &= 4(p_\xi p_x + p_\eta p_y + p_\zeta p_z). \end{aligned}\right\} \quad . \quad . \quad (9)$$

Since the equation (5) does not contain explicitly the variables ξ, η, ζ, the distance r depending only on x, y, z, the differential coefficients p_ξ, p_η, p_ζ may be regarded as constants. The physical meaning of this is, as is shown by (8) and (2), that the total momentum of the system is constant during the motion.

Substituting (9) in (5), we obtain a quadratic differential equation in the three variables x, y, z. This equation describes the relative motion of the two masses.

We shall confine our investigation to the case when the total momentum is equal to zero, and put accordingly

$$p_\xi = p_\eta = p_\zeta = 0. \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Then the second and third terms of equation (5) dis-

appear, and the result is

$$\Sigma p_x^2 = \frac{1}{4c^2}\left(\alpha_1' - \frac{e_1 e_2}{r}\right)^2 - \tfrac{1}{2}(m_1^2 + m_2^2)c^2 + \frac{c^6(m_1^2 - m_2^2)^2}{4\left(\alpha_1' - \frac{e_1 e_2}{r}\right)^2}. \quad (11)$$

It is easily seen that the relative motion is a plane one. Therefore we can describe the motion by using polar coordinates in the plane of relative motion. We can do that, for instance, by choosing the z-axis perpendicular to the plane of motion and by putting further

$$x = r\cos\phi, \qquad y = r\sin\phi.$$

Then we have

$$\Sigma p_x^2 = p_r^2 + \frac{p_\phi^2}{r^2}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

where p_r and p_ϕ are the partial differential coefficients of the function W corresponding to the variables r and ϕ.

Substituting (10) into (11), we see that our equation does not contain explicitly ϕ, and that accordingly p_ϕ may be considered as a constant. This means physically that the relative angular momentum is constant throughout the motion. We put

$$p_\phi = \alpha_2.$$

Equation (11) becomes

$$p_r^2 = \frac{1}{4c^2}\left(\alpha_1' - \frac{e_1 e_2}{r}\right)^2 - \tfrac{1}{2}(m_1^2 + m_2^2)c^2 + \frac{c^6(m_1^2 - m_2^2)}{4\left(\alpha_1' - \frac{e_1 e_2}{r}\right)^2} - \frac{\alpha_2^2}{r^2}.$$

. . . (13)

Having thus found the value of p_r in terms of r only, and knowing p_ϕ to be constant, we can write the solution of our equation.

For brevity's sake we shall write $\mathrm{P}(1/r)$ for the right side of (13). Then the solution is

$$\mathrm{W} = \int \sqrt{\mathrm{P}(1/r)}\, dr + \alpha_2\phi. \quad . \quad . \quad . \quad . \quad (14)$$

The equation of the orbit being

$$\frac{\partial W}{\partial \alpha_2} = \beta_2,$$

where β_2 is another constant, we obtain from (14) the equation of the orbit in relative motion in the form

$$\phi - \beta_2 = -\int \frac{\alpha_3 d(1/r)}{\sqrt{P(1/r)}}, \quad \ldots \ldots \quad (15)$$

which we have now to discuss. We shall choose the particular values of β_2 so as to make an apex of the curve, if there is any, to correspond to $\phi = 0$.

4. We shall put

$$\alpha_1' = \alpha_1 + (m_1 + m_2) e^2,$$

where α_1 is a constant expressing the sum of the kinetic and potential energies of our system.

If we confine our attention to cases which are of greater practical interest, we may consider the kinetic energy as small compared with the energy accumulated in the masses when at rest. This comes to supposing that

$$\frac{\alpha_1 - \frac{e_1 e_2}{r}}{(m_1 + m_2) c^2} \quad \ldots \ldots \quad (16)$$

is small compared with unity. If the expression (13) for $P(1/r)$ is developed in powers of (16), and the powers higher than the second are neglected, then it assumes the form

$$P(1/r) = \frac{2 m_1 m_2}{m_1 + m_2}\left(\alpha_1 - \frac{e_1 e_2}{r}\right) + \frac{1}{c^2}\left[1 - \frac{3 m_1 m_2}{(m_1 + m_2)^2}\right]\left(\alpha_1 - \frac{e_1 e_2}{r}\right)^2 - \frac{\alpha_2^2}{r^2} \quad \ldots \quad (17)$$

This expression being quadratic in $1/r$, the integral in (15) is of elementary kind, and the form of the orbit can be easily discussed.

If we put

$$\left.\begin{aligned}
\eta^2-1 &= \frac{2\alpha_1(m_1+m_2)\left[1+\frac{\alpha_1}{2c^2}\left(\frac{m_1^2-m_1m_2+m_2^2}{m_1^2m_2+m_1m_2^2}\right)\right]\left\{\alpha_2^2-\left(\frac{e_1e_2}{c}\right)^2\left[1-\frac{3m_1m_2}{(m_1+m_2)^2}\right]\right\}}{m_1m_2e_1^2e_2^2\left[1+\frac{\alpha_1}{c^2}\left(\frac{m_1^2-m_1m_2+m_2^2}{m_1^2m_2+m_1m_2^2}\right)\right]^2}, \\
p &= -\frac{\alpha_2^2-\left(\frac{e_1e_2}{c}\right)^2\left[1-\frac{3m_1m_2}{(m_1+m_2)^2}\right]}{\frac{m_1m_2e_1e_2}{m_1+m_2}\left[1+\frac{\alpha_1}{c^2}\left(\frac{m_1^2-m_1m_2+m_2^2}{m_1^2m_2+m_1m_2^2}\right)\right]}, \\
\gamma^2 &= 1-\left(\frac{e_1e_2}{\alpha_2c}\right)^2\left[1-\frac{3m_1m_2}{(m_1+m_2)^2}\right],
\end{aligned}\right\} \quad (18)$$

then for positive values of γ^2 equation (15) will take the form

$$\frac{p}{r} = 1+\eta\cos(\gamma\phi). \quad . \quad . \quad . \quad . \quad . \quad (19)$$

If, however, γ^2 is negative, we shall have to put

$$\gamma = i\delta,$$

and shall obtain instead of (19)

$$\frac{p}{r} = 1-\eta\cosh(\delta\phi). \quad . \quad . \quad . \quad . \quad . \quad (20)$$

The intermediate case $\gamma^2=0$ leads to an algebraic quadratic equation between r and ϕ, which need not be written down, because the properties of the corresponding orbit as the intermediate orbit between the orbits (19) and (20) are sufficiently obvious.

We see that the equations (19) and (20) are exactly the same as in the case of a fixed centre of force. The whole difference lies in the values of the three constants η, p, and γ, which are modified by our supposition that both masses are moving under their mutual attraction or repulsion. We obtain the formulæ for a fixed mass m_2 by putting $\frac{m_2}{m_1}=\infty$.

If, on the other hand, we put in (18) $c=\infty$, we obtain the formulæ of the classic problem of two bodies. In this case we have $\gamma=1$, and the relative trajectories are conics.

It is hardly necessary to discuss the equations (19) and (20) in full detail.

The curve (19) may be described as generated by the superposition of a motion in a conic and a rotation of the perihelion. The constant γ^2 being smaller than unity, the rotation is in the same direction in which the mass moves along the conic.

The curve (20) is in most cases a spiral (and so is the intermediate curve $\gamma=0$). If $\eta>1$, both ends of the curve are spirals winding round the origin. This can obviously happen only in the case of attraction, as it requires that α_1 should be negative. The value of p in (20) being negative, the curve is real for all values of ϕ between $-\infty$ and $+\infty$. But if $\eta<1$, only one end of the curve is a spiral if e_1 and e_2 are of different sign, the other one stretching to infinity. In this case α_1 must be positive, which does not lay any restriction on the signs of the charges. If there is attraction, p is negative and the curve is included between the asymptote $\phi=\pm\frac{1}{\delta}\operatorname{arccosh}\frac{1}{\eta}$ and $\phi=\pm\infty$, where the signs have to be taken either both positive or both negative. In the case of repulsion, p is positive and the curve is included between the two asymptotes $\phi=\frac{1}{\delta}\operatorname{arccosh}\frac{1}{\eta}$. This time it is not a spiral, but a distorted branch of an hyperbola.

The critical value of the angular momentum corresponding to $\gamma=0$ is in our case

$$\alpha_2^2=\left(\frac{e_1e_2}{c}\right)^2\left[1-\frac{3m_1m_2}{(m_1+m_2)^2}\right] \quad . \quad . \quad . \quad (21)$$

compared with

$$\alpha_2^2=\left(\frac{e_1e_2}{c}\right)^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (22)$$

in the case of a fixed centre. Thus for a moving centre the critical value is smaller than for a fixed one. The effect is largest for $m_1=m_2$, (21) giving then a twice smaller value than (22). For a hydrogen atom the effect is a negligible one, roughly 1 in 1200; for a helium ion, nearly 1 in 5000.

We come to the conclusion that on the theory of relativity the relative motion of two charges is of much the same kind as the motion round a fixed centre of attraction. In cases of practical importance the quantitative difference between them is only small.

But we must not forget that the conclusions of this paragraph are only valid so long as the kinetic energy is small compared with the energy stored in the masses. This means that our equations give an approximately accurate description of those parts of the orbits where r is not too small. Thus equation (20) describes accurately only the outer parts of spiral trajectories. Knowing nothing about the inner part of it, we cannot even decide whether for given value of α_2 the real orbit will be a spiral one or not. The critical value (21) determines only the analytical form of the equation for the outer part, but, as we shall see in the next paragraph, does not provide a criterion which would allow us to decide whether the orbit is a spiral one or not.

5. The investigation of the last paragraph was based on the assumption that the sum of the kinetic and potential energies of the system is small compared with the energy stored in the masses.

The solution of the problem for the exact value (13) of $P(1/r)$ is far more difficult. But if we do not attempt to discuss the form of the trajectories in full detail, but limit our attention to the question whether spiral trajectories still occur and under what condition, the problem is a very simple one, and the form of $P(1/r)$ provides us all the information that we want. According to (13), $P(1/r)$ must be positive for any real motion. Spiral trajectories will accordingly be possible whenever the right side of (13) is positive for infinitely decreasing values of r. In other words, we may have spiral trajectories if the coefficient of $1/r^2$ is positive, but not if it is negative. This leads again to a critical value for α_2. The coefficient of $1/r^2$ being $\left(\frac{e_1e_2}{2c}\right)^2 - \alpha_2^2$, the condition for spiral trajectories is

$$\alpha_2^2 < \left(\frac{e_1e_2}{2c}\right)^2. \quad . \quad . \quad . \quad . \quad . \quad . \quad (23)$$

This is the value which we obtained in the previous paragraph for the case of equal masses, using an approximate method.

The physical meaning of the constant α_2 may not be obvious from the above formulæ. We can easily show that it really is the quantity which must be called the angular momentum of relative motion on the Einstein theory. It follows from (7), (10), and (12),

$$\Sigma p_{x_1}^2 = \Sigma p_{x_2}^2 = \Sigma p_x^2 = p_r^2 + \frac{p_\phi^2}{r^2}.$$

Substituting this into (4), we have one of the canonical differential equations :

$$\dot{\phi} = \frac{\partial H}{\partial p_\phi} = \frac{c p_\phi}{r^2}\left[\frac{1}{\sqrt{m_1^2 c^2 + \Sigma p_{x_1}^2}} + \frac{1}{\sqrt{m_2^2 c^2 + \Sigma p_{x_2}^2}}\right].$$

Putting $p_\phi = \alpha_2$, and using (2), we find

$$\alpha_2 = \frac{r^2 \dot{\phi}}{\frac{\sqrt{1-\beta_1^2}}{m_1} + \frac{\sqrt{1-\beta_2^2}}{m_2}}.$$

As in the classical problem of two bodies, the rôle of the reciprocal value of the mass is played by the sum of the reciprocal values of the two masses, both expressed in terms of their velocities. Thus α_2 really deserves the name of the angular momentum of relative motion.

6. The results obtained in the previous pages may be summed up as follows :—

The Hamilton-Jacobi equation for two charges moving under mutual attraction or repulsion with masses varying according to Einstein's law can be easily solved, when the total momentum of the system is equal zero.

The exact solution of the equation shows that, if there is attraction, for small values of relative angular momentum the relative trajectory is a spiral, as in the case of fixed centre of attraction. The critical value of angular momentum is half of its value in the latter case.

The discussion of the exact solution in detail is not attempted. For practical purposes it is sufficient to discuss the approximate solution, which is obtained by supposing the kinetic energy of the system small compared with the energy stored in the masses. On this assumption the explicit equations of the trajectories are the same as for a fixed centre. Only the values of the constants in these equations are modified. The effect on the hydrogen or helium spectrum brought about by this modification is, however, so small that it could not be observed.

II. *Study of the mutual Magnetic Action of two Moving Charges.*

7. There is another problem which is in many ways related to the one studied above, and which can easily be solved by the same method. The atomic theory in its

simplest form takes no account of the magnetic forces arising from the mutual displacement of charges in the atom ; in fact, these forces in an atom of hydrogen or in a helium ion are so small compared with electrostatic ones that they have to be neglected.

In atoms containing more than one electron these forces probably cannot be neglected. D. Roschdestvensky and G. Krutkow *, and independently myself, some two years ago came to the conclusion that these forces account for the doublets in the arc spectra of alkali metals. A. Sommerfeld †, in the 3rd edition of his excellent book 'Atombau und Spektrallinien,' gives a theory of doublets and triplets based on the same idea. But these are not the facts with which we propose to deal in the present paper.

As in the previous section, we are going to study the movement of a system consisting of only two charges ; so that our result will bear only on the theory of the hydrogen atom and the helium ion. We intend to find out in what way the relative trajectories of two charges are modified by the magnetic field arising from their motion. We shall be able to compute the influence of these forces on the spectra emitted by such systems. However small this influence may be, it is not uninteresting from a theoretical point of view.

8. We must confine our attention to cases of *quasi-stationary* movement of the charges. This means that the changes in the state of the system are supposed to be so slow that we may neglect the time necessary for the propagation of the electromagnetic disturbance from one charge to the other. In atomic theory we are certainly entitled to do so. Then the Lagrangean function of our system has the form

$$\mathrm{L} = \frac{m_1 v_1^2}{2} + \frac{m_2 v_2^2}{2} + \gamma(\dot{x}_1\dot{x}_2 + \dot{y}_1\dot{y}_2 + \dot{z}_1\dot{z}_2) - c^2\gamma, \quad . \quad (24)$$

where we put for brevity's sake

$$\frac{e_1 e_2}{c^2 r} = \gamma, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (25)$$

* D. Roschdestvensky, 'Verhandlungen des Optischen Instituts in Petersburg,' Berlin, 1921.

† A. Sommerfeld, 'Atombau und Spektrallinien,' Braunschweig, 1922.

The third term on the right side of (24) accounts for magnetic forces. We consider the masses as constant.

The momenta of the masses are :

$$\left.\begin{aligned} p_{x_1} &= \frac{\partial L}{\partial \dot{x}_1} = m_1\dot{x}_1 + \gamma\dot{x}_2, \\ p_{x_2} &= \frac{\partial L}{\partial \dot{x}_2} = \gamma\dot{x}_1 + m_2\dot{x}_2, \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (26)$$

and similarly for the two other axes.

The Hamiltonian function as computed by (3) differs from the Lagrangean function (24) only by the sign of the last term. Solving (26), we find

$$\dot{x}_1 = \frac{m_2 p_{x_1} - \gamma p_{x_2}}{m_1 m_2 - \gamma^2},$$

$$\dot{x}_2 = \frac{m_1 p_{x_2} - \gamma p_{x_1}}{m_1 m_2 - \gamma^2}.$$

Substituting this and the analogous expressions of the remaining four components of velocities in the formula for the Hamiltonian function, we find

$$H = \frac{1}{m_1 m_2 - \gamma^2}\left\{ \frac{m_2}{2}(p_{x_1}^2 + p_{y_1}^2 + p_{z_1}^2) + \frac{m_1}{2}(p_{x_2}^2 + p_{y_2}^2 + p_{z_2}^2) - \gamma(p_{x_1}p_{x_2} + p_{y_1}p_{y_2} + p_{z_1}p_{z_2})\right\} + c^2\gamma. \quad (27)$$

Putting the right side equal to the total energy α_1, we obtain the differential equation of Hamilton-Jacobi.

9. We proceed to solve this equation by the same method as we used in § 3. Using the variables (6) and the formulæ (7), we bring the equation to the form

$$\frac{1}{m_1 m_2 - \gamma^2}\left\{ \left(\frac{m_1 + m_2}{2} - \gamma\right)\Sigma p_\xi^2 + \left(\frac{m_1 + m_2}{2} + \gamma\right)\Sigma p_x^2 + (m_2 - m_1)\Sigma p_x p_\xi \right\} + c^2\gamma = \alpha_1. \quad . \quad (28)$$

We see again that the total momentum whose components are p_ξ, p_η, p_ζ is constant throughout the motion.

If, again, we limit our attention to the case

$$p_\xi = p_\eta = p_\zeta = 0,$$

and introduce polar coordinates in the plane of relative motion, (28) takes the simpler form

$$\left(\frac{m_1+m_2}{2}+\gamma\right)\left(p_r^2-\frac{p_\phi^2}{r^2}\right)=(\alpha_1-c^2\gamma)(m_1m_2-\gamma^2). \quad (29)$$

For practical purposes we may consider γ as small compared with m_1 or m_2. Indeed, we see from (25) that the absolute value of γ is the amount of mass which on the theory of relativity corresponds to the mutual potential energy of the charges. This amount is always small compared with the masses themselves. Thus we can further simplify (29) and obtain the equation

$$p_r^2=\frac{2m_1m_2}{m_1+m_2}\left(\alpha_1-\frac{e_1e_2}{r}\right)\left(1-\frac{2e_1e_2}{(m_1+m_2)c^2}\cdot\frac{1}{r}\right)-\frac{\alpha_2^2}{r^2}. \quad (30)$$

Here we have substituted for γ its value (25) and put $p_\phi=\alpha_2$. Putting $S(1/r)$ for the right side of (30), we have the solution of the differential equation of Hamilton-Jacobi :

$$W=\int\sqrt{S(1/r)}\,dr+\alpha_2\phi. \quad . \quad . \quad . \quad (31)$$

The equation of the orbit in relative motion is

$$\beta_2=\alpha_2\int\frac{d(1/r)}{\sqrt{S(1/r)}}+\phi. \quad . \quad . \quad . \quad . \quad (32)$$

Since $S(1/r)$ is again a quadratic function of $1/r$, the explicit form of the equation of the trajectory is again (19) or (20) if the value of β_2 in (32) is chosen so as to make an apex of the curve to correspond to $\phi=0$. The constants in these equations have the values

$$\left.\begin{aligned}\eta^2-1&=\frac{2\alpha_1(m_1+m_2)\left[\alpha_2^2-\dfrac{4m_1m_2e_1^2e_2^2}{(m_1+m_2)^2c^2}\right]}{m_1m_2e_1^2e_2^2\left[1+\dfrac{2\alpha_1}{(m_1+m_2)c^2}\right]^2},\\ p&=-\frac{\alpha_2^2-\dfrac{4m_1m_2e_1^2e_2^2}{(m_1+m_2)^2c^2}}{\dfrac{m_1m_2}{m_1+m_2}e_1e_2\left[1+\dfrac{2\alpha_1}{(m_1+m^2)c^2}\right]},\\ \gamma&=\sqrt{1-\frac{4m_1m_2e_1^2e_2^2}{\alpha_2^2(m_1+m_2)^2c^2}}.\end{aligned}\right\} \quad (33)$$

But again, as in the former case, we must keep in mind that our formulæ apply only to orbits or to parts of orbits where the two charges do not come too close together. So if for a given very small value of α_2 a part of the orbit can be with sufficient accuracy described by equation (20), we cannot conclude that this orbit will be a spiral throughout. The accurate equation (29) shows rather that spiral trajectories in this case are impossible. But we need not dwell on these details.

Formulæ (33) describe the disturbance of the motion caused by the magnetic field which it produces. They show that, for an originally elliptic orbit, both the greater semi-axis and the excentricity are slightly altered by the magnetic forces. And, what is more important, the orbit is no longer a closed curve, but there is a slow progression of apses, because γ differs from unity. All these disturbances are very small. It is easily seen from (33) that they are largest for $m_1 = m_2$ and disappear as well for $\frac{m_1}{m_2} = 0$ as for $\frac{m_1}{m_2} = \infty$. This result could be easily anticipated. Indeed, if one of the masses is large compared with the other, its velocity is small, and so is the magnetic field produced by its motion. The order of magnitude of the disturbances is the same as the order of the corrections to the relativity effect found in the first part of this paper. If m_1 is the mass of an electron and m_2 the mass of the nucleus of a hydrogen or helium atom, and we call x the order of the disturbance produced by the relativity variation of mass, then the order of the disturbances under consideration is $\frac{m_1 x}{m_2}$.

It is interesting to compare the last formulæ of the sets (18) and (33). If we measure by $1-\gamma^2$ the rotation of apses, we see from (18) that the rotation produced by the relativity effect is decreased by $\frac{3m_1m_2}{(m_1+m_2)^2}\left(\frac{e_1e_2}{\alpha_2 c}\right)^2$ when we take into account the motion of both masses. On the other hand, (33) shows that the effect of the magnetic forces is to increase it by $\frac{4}{3}$ of the same quantity.

10. From what has been stated already, we gather that the magnetic forces must cause a structure of spectral lines of the same kind as is produced by the relativity effect,

only on a scale $\frac{m_1}{m_2}$ times smaller. This ratio for a hydrogen atom being 1/1840, the effect cannot be observed experimentally. We give a brief computation of its value. The quantum equations determining the values of α_1 and α_2 are:

$$2\int_{r_1}^{r_2} \sqrt{S(1/r)}\, dr = n_1 h,$$

$$\int_0^{2\pi} \alpha_2 d\phi \qquad = n_2 h,$$

where r_1 and r_2 are the roots of the equation $S(1/r)=0$. This gives

$$\left.\begin{aligned} -2\pi\Bigg\{ e_1 e_2 \sqrt{\frac{m_1 m_2}{-2\alpha_1(m_1+m_2)}}\left[1+\frac{2\alpha_1}{(m_1+m_2)c^2}\right] \\ +\sqrt{\alpha_2^2 - \frac{4m_1 m_2 e_1^2 e_2^2}{(m_1+m_2)^2 c^2}}\Bigg\} &= n_1 h, \\ 2\pi\alpha_2 &= n_2 h. \end{aligned}\right\} \quad (34)$$

The constant α_1 measures the sum of kinetic and potential energy of the system, and is very small compared with $\epsilon=(m_1+m_2)c^2$, which is the amount of energy stored in the masses. Thus the second term in the square bracket of the first equation (34) is very small compared with unity. And so is the expression $\frac{4m_1 m_2 e_1^2 e_2^2}{\alpha_2^2 (m_1+m_2)^2 c^2}$. If we neglect them, we find in first approximation for α_1 the value

$$-\alpha_1^0 = \frac{2m_1 m_2 \pi^2 e_1^2 e_2^2}{(m_1+m_2)\, h^2 (n_1+n_2)^2}. \quad . \quad . \quad . \quad . \quad (35)$$

This is the value of α_1 when magnetic forces are neglected. Substituting this approximate value in the small term of (34), and neglecting the powers of $\frac{\alpha_1^0}{\epsilon}$ which are higher than the first, we find easily

$$-\alpha_1 = -\alpha_1^0\left[1-\frac{8\alpha_1^0}{\epsilon}\left(\frac{1}{2}+\frac{n_1}{n_2}\right)\right]. \quad . \quad . \quad . \quad (36)$$

The corresponding formula for the relativity effect * is

$$-\alpha_1 = -\alpha_1^0\left[1-\frac{2\alpha_1^0}{\epsilon'}\left(\frac{1}{4}+\frac{n_1}{n_2}\right)\right], \quad . \quad . \quad . \quad (37)$$

where $\epsilon'=m_1c^2$ is the amount of energy stored in the electron. We see that both effects are of much the same kind: the spectral lines are shifted and split into a number of components. But the relativity effect is by much the larger, since we have approximately $\epsilon'=\frac{m_1\epsilon}{m_2}$. Thus the shifting effect of the magnetic action is $\frac{8m_1}{m_2}$ times that of the relativity, and the splitting effect $\frac{4m_1}{m_2}$ times. For a hydrogen atom the figures are 1/230 and 1/460.

For the Balmer series of hydrogen, the end trajectories correspond to $n_1+n_2=2$: a circular one with $n_1=0$, $n_2=2$, and an elliptic one with $n_1=n_2=1$. On account of the relativity effect, the lines are doublets for which the theoretical value of difference of wave-numbers is 0·365. The last decimal of this figure is already beyond the limit of accuracy of modern wave-length measurements. Accordingly the 430 times smaller effect of the magnetic forces can in no way be checked experimentally.

The effect of magnetic forces is to increase slightly the relativity effect. But from the formulæ of the first part of this paper we can easily show that the relativity effect is on the other hand slightly decreased by the fact that the nucleus is not fixed, as it is assumed to be in Sommerfeld's theory. The correction is of the same order as the one just computed. Thus both corrections partly cancel each other.

* See, for instance, J. H. Jeans, 'The Dynamical Theory of Gases,' 3rd edition, p. 424.

XI. *On the Theory of Metallic Conduction.* *By* Tch. Bialobjeski, *Professor in the University of Cracow (Poland)* *.

THE electron theory of metallic conduction, notwithstanding some serious objections raised against it, is still in favour at the present time as giving the best interpretation of facts in this field of physical inquiry. But the investigations of Kamerlingh Onnes on the electrical conductivity of some pure metals in the proximity of absolute zero (0° K.) showed that insurmountable difficulties stand here in the way. Kamerlingh Onnes discovered that at the temperature of liquid helium (about 4° K.) some metals pass into a state of super-conductivity, their specific resistance being less than one 10^{-11} part of that at 0° C. The superconducting state appears abruptly at a definite temperature. On the free-electron theory the electrical conductivity of a metallic body is given by the formula

$$c = \frac{ne^2l\bar{v}}{6\mathrm{RT}}\ \dagger,$$

where n denotes the number of free electrons in unit volume, e is the elementary charge of electricity, l the mean free path of an electron, $\bar{v}$ the mean velocity of thermal agitation of electrons, R the gas constant calculated for one molecule, T the absolute temperature. An enormous increase of conductivity at 0° K. might be attributed to an increase of the free path; but such an almost unlimited lengthening is hardly conceivable. Other attempts aiming at reconciliation of the theory of free electrons with experimental data cannot be regarded as satisfactory.

In the year 1915 Sir J. J. Thomson restated his theory which had been considered in the book on the 'Corpuscular Theory of Matter' as the second theory of metallic conduction, the first being that of free electrons ‡.

According to this theory atoms of metals contain electrical doublets, *i. e.* pairs of equal and opposite charges at short distance one from the other. In the absence of external electrical force the axes of doublets are uniformly distributed, having all directions in space. When an electrical force acts on the body the axes of the doublets tend to point

* Communicated by the Author.

† See, for instance, Jeans, 'Dynamical Theory of Gases,' 2nd edition, p. 319.

‡ Phil. Mag. xxx. p. 192 (1915).

in the direction of the force and the moments of the doublets acquire a finite resultant; whose direction of course coincides with that of the force. If no disturbing influences were at work, the smallest electrical force would be sufficient to produce the maximum polarization, *i. e.* to arrange the axes of all doublets in one direction. The thermal agitation (and perhaps some other influences) put a limit to the polarization produced by a given electric force.

A characteristic feature of the new aspect of Sir J. J. Thomson's theory is the introduction of an intermolecular field of force; this makes the view of the conductivity of metals similar to that by which Weiss explained ferromagnetism. It is assumed that, when the external force is applied, a new force arises acting on the doublets and proportional to the polarization of the body. If N is the number of doublets in unit volume, M the electrical moment of a doublet, I the polarization of unit volume (that is, the resultant of all moments in the direction of electric force), we might imagine that $\frac{I}{M}$ doublets have their axes pointing in the direction of the force and forming atomic chains, while the axes of the other doublets are distributed uniformly in all directions. The peculiarity of metals among other substances is that some electrons are very loosely connected with the atoms. Consequently along atomic chains an interchange of electrons between atoms is going on continuously. Let us suppose that p electrons pass along each chain per second. Denoting by d the distance of the adjacent atoms in the chain, we have $\frac{Id}{M}$ chains crossing unit area at right angles to the electric force. The current through unit area is then

$$i = ep\frac{Id}{M} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The relation between I and the extraneous electric force is very complicated; indeed, the disturbing effect of molecular motion and the action of intermolecular forces must be taken into account. The explanation of super-conductivity, however, presents no difficulty. The rôle of electric force in metallic conduction is mainly to form chains; at low temperatures the disturbing effect of thermal energy vanishes, so that intermolecular forces become preponderant: the chains persist after the electric force is removed and electrons are transmitted continually along them.

The form given to the theory by Sir J. J. Thomson seems to be open to objection. The recent progress of our views on the dependence of various phenomena upon temperature has led to the conviction that Planck's quanta should be introduced into all problems concerning partition of energy. As the thermal agitation plays a prominent part both in electrical and thermal conductivity, the introduction of quanta into the theory of these phenomena seems unavoidable. Therefore Thomson's theory, resting on the rule of equipartition and taking no account of striking departures from this rule at the proximity of 0° K., must be regarded as insufficient. Moreover, the assumption of intermolecular fields of force is of a somewhat vague character. It is the object of this paper to show that the hypothesis of intermolecular fields of force, considered as an important agent of conductivity at low temperatures, becomes unnecessary if we introduce energy quanta. It will be seen that combining Sir J. J. Thomson's hypothesis regarding the interchange of electrons between atoms with the fundamental relations of the quantum-theory, we are led to an interpretation of essential features shown by the conduction of electricity through metals.

Let X stand for the extraneous electric force acting on a metallic conductor. Applying the well-known formula of Boltzmann, Langevin has obtained in his theory of paramagnetism an expression for the polarization of unit volume, which may be used in our case without any change. We have

$$I = NM\left(\frac{e^x + e^{-x}}{e^x - e^{-x}} - \frac{1}{x}\right),$$

where $x = \frac{MX}{RT}$. According to classical statistical theory, $\frac{1}{2}RT$ represents the mean energy corresponding to one degree of freedom in a system, which is in thermal equilibrium with a gas at T° K. At present the opinion has prevailed that this statistical inference is only true at high temperatures and is utterly inapplicable at very low temperatures. An expression for the mean energy belonging to one degree of freedom has been given by Planck in the first form of his theory as

$$\frac{1}{2}\frac{h\nu}{e^{\frac{h\nu}{RT}} - 1},$$

h being Planck's universal constant, ν the frequency of

oscillations corresponding to the degree of freedom under examination. We assume the thermal agitation to be resolved as an assemblage of homogeneous vibrations having a definite scale of frequencies. It is known that Debye, and independently Born and Kárman, have analysed in this way the thermal vibrations of bodies with crystalline structure. It will be natural to suppose that the atoms of metals containing doublets are performing similar vibrations, by which the directive action of electric force is disturbed in accordance with Boltzmann's rule giving Langevin's formula. Consequently we retain unchanged the expression of I putting herein

$$x = \mathrm{MX} : \frac{h\nu}{e^{\frac{h\nu}{\mathrm{RT}}} - 1}.$$

In order to obtain simple relations let us admit only one frequency of atomic vibrations. Although this assumption might be deemed a too great simplification, the general features of the phenomena under consideration will not be essentially modified. The electric current will now be represented by the expression :

$$i = \frac{epd\mathrm{I}}{\mathrm{M}} = \mathrm{N}epd\left(\coth x - \frac{1}{x}\right), \quad . \quad . \quad . \quad . \quad (2)$$

where $x = \dfrac{\mathrm{MX}\left(e^{\frac{h\nu}{\mathrm{RT}}} - 1\right)}{h\nu}$.

Let us call Langevin's function and denote by $\mathrm{L}(x)$ the expression in brackets. Expansion of $\mathrm{L}(x)$ in an ascending series gives

$$\mathrm{L}(x) = \frac{x}{3} - \frac{x^3}{45} + \ldots .$$

At high temperatures we may use approximate formulas:

$$x = \frac{\mathrm{XM}}{\mathrm{RT}} \quad \text{and} \quad \mathrm{L}(x) = \frac{x}{3} = \frac{\mathrm{XM}}{3\mathrm{RT}}.$$

In this case

$$i = \frac{\mathrm{NM}epd}{3\mathrm{RT}} \cdot \mathrm{X}; \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

the conductivity at high temperatures will be expressed by

$$c = \frac{\mathrm{NM}epd}{3\mathrm{RT}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (3')$$

More precise results can be obtained by using the results of Debye's calculation relating to the spectrum of thermal vibrations *. When N molecules participate in the elastic vibrations (identified with the thermal energy of the body), then the number of vibrations with frequencies contained between ν and $\nu+d\nu$ is

$$3N\frac{\nu^2 d\nu}{\nu_m^3};$$

the maximum frequency of the vibrations is here denoted by ν_m. Thus

$$i=\frac{3Nepd}{\nu_m^3}\int_0^{\nu_m} L(x)\nu^2 d\nu. \quad . \quad . \quad . \quad . \quad (4)$$

We do not endeavour to investigate this equation exactly and shall content ourselves with the examination of conductivity at higher temperatures. Writing approximately

$$e^x=1+x+\tfrac{1}{2}x^2,$$

we get

$$L(x)=\tfrac{1}{3}XM\left(\frac{1}{RT}+\frac{h\nu}{2R^2T^2}\right),$$

and by integration

$$i=\frac{NMepd}{3RT}\left(1+\tfrac{3}{8}\frac{\theta}{T}\right)X, \quad . \quad . \quad . \quad . \quad (5)$$

introducing Debye's characteristic temperature by the equality

$$\theta=\frac{h\nu_m}{R}.$$

Thus the temperature coefficient of conductivity is greater than that of perfect gases. This conclusion is in accordance with experimental data concerning pure metals. Iron and nickel on Debye's theory of specific heat are distinguished by an exceptionally high value of the characteristic temperature ($\theta^{Fe}=467$, $\theta^{Ni}=435$); the same metals have particularly considerable temperature coefficients of conductivity which agrees with (5).

Using the simplest expression (3′) for conductivity at comparatively high temperatures we proceed to calculate

* *Annalen d. Physik*, xxxix. p. 789 (1912).

the value of p, the number of electrons which pass per second along the atomic chain :

$$p = \frac{3\mathrm{RT}c}{\mathrm{NM}ed}.$$

We will apply this equation to the case of silver ($c = \frac{1}{1500}$ E.M.U. at 0° C.). For the values of constants R, N, e we take :

$$\mathrm{N} = \frac{6{\cdot}1 \,.\, 10^{23} \,.\, 10{\cdot}5}{108}, \quad e = 1{\cdot}6 \,.\, 10^{-20}, \quad \mathrm{R} = 1{\cdot}4 \,.\, 10^{-16}.$$

The moment of a doublet $\mathrm{M} = e\delta$, where δ is the distance of two oppositely charged constituents, we take to be equal to $1{\cdot}6 \,.\, 10^{-20} \,.\, 10^{-8}$. For the distance of two neighbouring atoms we put $d = 3 \,.\, 10^{-8}$ (equal to the linear constant of the crystalline lattice of calcite). The two last numbers do not pretend to give more than the true order of magnitude. We obtain thus $p = 1{\cdot}7 \,.\, 10^{16}$. It will be observed that this number corresponds to frequencies of electromagnetic vibrations in the region lying between the ultraviolet and the softest X radiations.

From what has been said it follows that the crucial test of the theory will be its application to the conductivity at very low temperatures. We note that at 0° K., x is equal to exponential infinity and $\mathrm{L}(\infty) = 1$. Then

$$i_0 = \mathrm{N}epd, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

which gives for silver $4{\cdot}8 \,.\, 10^{11}$ E.M.U. This monstrous current arises if all the atoms of a piece of metal are arranged in parallel chains and may be started by an arbitrarily small electromotive force. It should not, however, be inferred that such a current can be excited almost spontaneously and instantly. In order to accumulate the enormous energy of magnetic field a corresponding amount of work must be done. As soon as the current has arisen, the smallest electric force would be sufficient for maintaining it : we have, therefore, super-conductivity. At the same time i_0 represents maximum or saturation current.

We shall now give the formulas expressing changes of current as a function of temperature and electromotive force. From (2) we deduce :

$$\frac{di}{d\mathrm{X}} = \frac{i_0}{\mathrm{X}}\left(\frac{1}{x} - \frac{4x}{(e^x - e^{-x})^2}\right), \qquad x = \frac{\mathrm{XM}}{h\nu}\left(\frac{h\nu}{e^{\overline{\mathrm{RT}}}} - 1\right). \quad (7)$$

At the proximity of 0° K., commencing, say, from 8° K., when x acquires very great values, the differential coefficient becomes exceedingly small; we have then a saturation current, independent of electromotive force and equal to i_0. At these temperatures the thermal agitation ceases to destroy perceptibly the directive action of electrical forces. We have also:

$$\frac{di}{dT} = i_0 \frac{XM}{RT^2} e^{\frac{h\nu}{RT}} \left(\frac{4}{(e^x - e^{-x})^2} - \frac{1}{x^2} \right). \quad . \quad . \quad (8)$$

The value of this derivative is also infinitely small near 0° K.

It seems that not all metals are able to pass into a superconducting state; some of them, at very low temperatures, tend to acquire a constant conductivity, enormous but finite. This behaviour is proper to metals containing foreign admixtures and especially to alloys. For the purpose of extending the theory to these cases, Sir J. J. Thomson introduces an additional restoring couple, which tends to make the axes of doublets to take up a definite direction, determined by the structure of the body.

This supposition may be immediately adapted to the form of theory here developed. Let us consequently admit that besides an electrical force and disturbing action of thermal energy there exists a couple D acting on the doublets. The equations (2) or (4) maintain their form, but the quantity denoted by x will be changed; we now have

$$x = XM : \left(\frac{h\nu}{e^{\frac{h\nu}{RT}} - 1} + D \right);$$

not only the thermal agitation, but also the couple D opposes the directive action of the electric force. It is evident that the limiting value of Langevin's function in the proximity 0° K., instead of unity will now be

$$\lim_{T=0} L(x) = \coth\left(\frac{XM}{D}\right) - \frac{XM}{D};$$

the current tends to the value

$$i_0' = Nepd\, L\left(\frac{XM}{D}\right). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

If D is not vanishingly small, i_0' will appear considerably less than i_0 and a clearly expressed saturation will not come

into existence. At higher temperature by using (4) and putting

$$L(x) = \frac{1}{3}\frac{XM}{RT+D}$$

we get

$$i = \frac{NMepd}{3(RT+D)}X;$$

therefore in the equation

$$\sigma = \sigma_0(1+\alpha_1 t)$$

giving the resistance σ as a function of temperature, the coefficient

$$\alpha_1 = \frac{\alpha}{1+\frac{\alpha D}{R}},$$

where α denotes the temperature coefficient of perfect gases. We see that $\alpha_1 < \alpha$, which is in agreement with experimental results in the case of alloys possessing probably a relatively strong restoring couple D.

We add in conclusion that by accepting the above deductions one must adopt a definite attitude with respect to the problem of existence of energy in a body at 0° K. In the second modified version, which Planck gave to his theory, a degree of freedom possesses a mean energy equal to

$$\frac{1}{2}\left(\frac{h\nu}{2}+\frac{h\nu}{e^{\frac{h\nu}{RT}}-1}\right).$$

If we had to put this value in the Langevin's function, we should obtain at 0° K.

$$x = \frac{2XM}{h\nu},$$

which in the case of not very high electromotive force is far less than unity. The super-conductivity could not be observed and even there would not exist a rapid increase of conductivity at low temperatures. We must therefore conclude that either metals do not possess energy at all at the absolute zero of temperature or the corresponding amount of vibratory energy $\frac{h\nu}{2}$ is not heat energy.

Cracow (Poland).
June 24, 1921.

XII. *The Evaluation of the Colours of the Spectrum in Terms of the Three Primary Colours. By* Dr. R. A. Houstoun, *Lecturer on Physical Optics in the University of Glasgow, and* Eric Dow, *B.Sc.* *

IF we take three colours—a red, a green, and a blue,—and represent them by the corners of an equilateral triangle, then any colour whatever can be represented by a point in the plane of the triangle. The three colours at the corners are referred to as the primary colours. Any colour obtained by mixing the primaries is represented by a point inside the triangle, and any colour, such as violet, that cannot be obtained by mixing the primaries, is represented by a point outside the triangle. Any colour obtained by mixing two colours is represented by a point on the straight line joining these two colours which divides the distance between them inversely in the ratio of their intensities. That is, to find the position of the mixture we regard the colours as particles the masses of which are proportional to their intensities, and proceed to find the centroid of these particles.

The laws of colour-mixing as expressed in the triangle diagram are facts of experience independent of all theories of colour-vision, and they are accepted by all theorists. It is the additional assumption, that to each of the primary colours there corresponds a primary-colour sensation, that has been a matter of dispute. The colour diagram was introduced by Newton in his 'Opticks,' and was fully developed by Helmholtz and Maxwell a century and a half later.

The position of the colours of the spectrum on the triangle diagram has been determined by Maxwell †, by König and Dieterici ‡, and by Abney § (twice); and the results of these observers have obtained considerable publicity, the curves of König and Dieterici in particular having been copied into many text-books. The different observers expressed their results in different units ; in 1919, Dr. Houstoun ‖ calculated them into the same units, and found that the agreement was not good, also that in no case were all the necessary particulars of the determination recorded. It was decided, therefore, to investigate the matter afresh.

Considerable care was bestowed on the selection of the

* Communicated by the Authors.
† Phil. Trans. 150, p. 57 (1861).
‡ *Berl. Ber.* p. 805, 1886.
§ Phil. Trans. 193A, p. 259 (1900), and 205A, p. 333 (1906).
‖ Phil. Mag. xxxviii. p. 402 (1919).

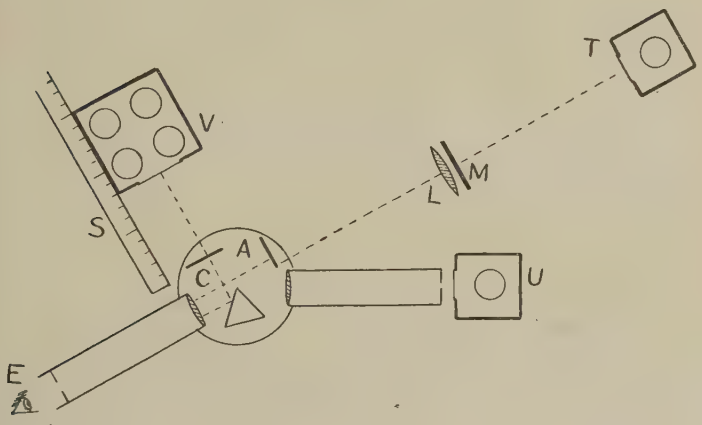
method. First of all a method based on Maxwell's colour box was tried. Thirty-two adjustable slits were set up at one side of a room in two rows of sixteen, and illuminated by tungsten-filament lamps; these were viewed through a prism and a telescope which had its eyepiece removed and a slit substituted. A biprism was mounted behind the prism. A preliminary determination by this apparatus was made by Mr. James Hyslop in the first months of 1920. The method was, however, abandoned, as it did not seem any better than Maxwell's.

The next attempt was based on König and Dieterici's method, and involved a quartz prism and nicol eyepiece, and was made in the autumn of 1920. This method was, however, little more than explored, as it seemed neither easy nor accurate. Finally, during the past winter a method was evolved which was both elegant and simple, and a great advance on the methods hitherto used. It does not make the determination easy; the actual matching of the colours will always remain a tedious business and require great care. But it removes all adventitious sources of error and worry connected with the apparatus, and leaves one with nothing but the actual difficulty of matching the colours itself. And as it does not require the intensities to be equal when the colours are matched, the matching is easier than by the other methods which involve heterochromatic photometry. Another advantage of the new method is the simple geometric procedure for evaluating the results.

Fig. 1 represents the apparatus. A is a ground-glass plate, called the comparison field. T is a 40-watt Osram lamp in a box with a square ground-glass window. Except for the window the box was light-tight, but at the same time adequate provision was made for its ventilation. The lens L focussed an image of the window on A. In front of the lens L was the colour mixer M; this consisted of a rectangular opening, across which one of three colour filters could be screwed. These filters were made of Wratten and Wainwright standard tricolour film mounted in B quality glass; the first was half red and half green, the second half green and half blue, and the third half blue and half red. If the first filter was screwed across the aperture, the film covering the latter changed from red to green, and the colour of the image on A changed simultaneously from red to green. For intermediate positions of the filter the colour of the image on A was a mixture of red and green, the amount of each component in the mixture being proportional to the

area it covered on the lens. The colour of the image on A was of course always perfectly uniform. The area of the lens covered by each colour could be read off on a scale. By means of the three filters the colour of the image on A could be changed from red through orange, yellow, and yellowish-green to green, from green through peacock-blue to blue, and from blue through purple, magenta, and carmine back to red.

Fig. 1.



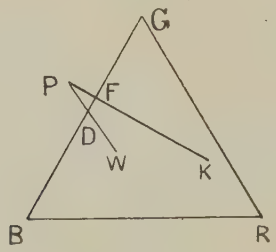
The ground-glass A stands on the table of a spectrometer, the telescope and collimator of which had been interchanged. In front of the collimator slit is a light-tight box U, with a ground-glass window containing a 40-watt Osram lamp. The prism is a 60° dense flint one, the faces of which measure $3{\cdot}5 \times 3{\cdot}1$ cm.^{2}; it stands at minimum deviation for Na light. The eyepiece of the telescope is removed and a slit substituted; this slit is 1 millimetre wide. An eye E looking in at the telescope consequently sees the face of the prism uniformly illuminated by whichever colour of the spectrum falls upon this slit and beside it the comparison field A. As the prism has a sharp edge, the two bound one another very sharply.

C is another ground-glass plate standing on the spectrometer table. It is illuminated by four 40-watt Osrams in a light-tight box V with a ground-glass window. The luminosity of C varies roughly according to the inverse square of its distance from the window. The eye at E sees the surface C superimposed by reflexion on the face of the

prism. The position of the box V can be read on the scale S.

Let us now consider the theory of the method. In fig. 2 the points R, G, and B represent the colours of the tricolour filters when placed in front of a diffusing surface illuminated by an Osram lamp. Consequently, when the colour mixer is fitted in succession with the red-green, the green-blue, and the blue-red filter, the comparison field can be given all the colours represented by the three sides RG, GB, and BR respectively. In the ordinary use of the apparatus, therefore, the comparison field is restricted to the sides of the triangle.

Fig. 2.



Suppose now that we wish to find the position on the diagram of a colour in the blue-green part of the spectrum. We know that the spectrum colours owing to their purity lie outside the triangle; the colour in question will lie at some such point as P. The collimator is turned until the face of the prism is illuminated with the colour in question. The lamps V (fig. 1) are then turned on; let us suppose that their light is represented by some such point as W. The face of the prism being illuminated by a mixture of white light and the spectrum colour is represented by a point on the line PW. The distances of the lamps V and the colour mixer are adjusted until the colours of the prism face and the comparison field are exactly the same. They must then both be represented by the point D. Now the ratio GD/DB is obtained from the reading of the colour mixer. Consequently the point D is fixed, and P is restricted to the line WD.

Let a piece of coloured gelatine be placed in front of C (fig. 1), so that the light obtained by reflexion from the prism has some such position as K (fig. 2). Then, as the box V is moved along the scale, the colour of the prism face

travels along the line KP. If the colour mixer and the distance of the box V are adjusted so that the face of the prism and the comparison field have exactly the same colour, this colour must be represented by the point F. The ratio GF/FB is then obtained from the reading of the colour mixer: consequently F is fixed, and P is restricted to the line KF. As it is now restricted to two intersecting straight lines, its position is completely determined. The positions of the other spectrum colours can be obtained in the same way. Instead of using the line KP we might have taken the line RP.

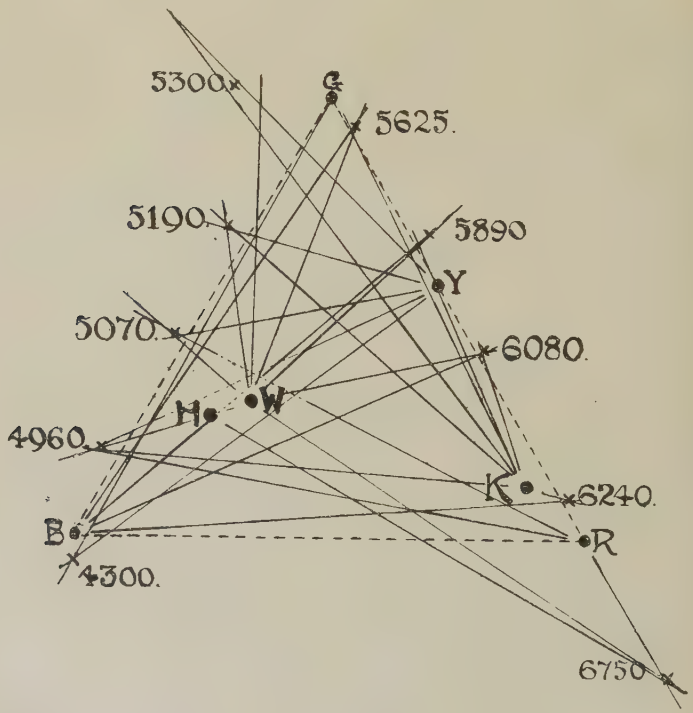
It remains to explain how the positions of W and K in fig. 2 were determined. Four such points were used altogether : W, which was obtained by letting the light from V fall directly on C, as already mentioned ; Y, which was obtained by placing a yellow filter (the Wratten minus-blue) immediately in front of C ; K, which was obtained by placing a pale red filter in front of C ; and H, which was obtained by substituting one 100-watt half-watt lamp for the four Osrams and not using any filter. Their positions on the diagram were obtained by turning out the lamp U (fig. 1), so that the prism face was illuminated solely by reflected light and manœuvring special filters each containing red, green, and blue film over the lens until a match was obtained. Then the areas of each colour covering the aperture were noted ; they specified the position of the point.

In matching the colours of the prism face and the comparison field, it is not necessary that their intensities should be exactly the same. It is, however, advisable that there should not be much difference between them. In addition to altering the distance of V and changing the width of the collimator slit, the intensities were also regulated by placing additional pieces of ground-glass in front of the windows in T and U, and by fixing wire gauze of different size of mesh immediately in front of M.

There were various screens for cutting off stray light which are not shown in fig. 1. There was a switch-board close to the observer's left hand for controlling the lamps T, U, V, and a reading-lamp, and beside the switch-board a cord for moving V, while with his right hand he could reach the screw motion for the colour mixer and the screw motion for the collimator. The spectrometer read to one minute of arc, and the lenses of telescope and collimator were of 1 inch diameter.

A series of results obtained by Mr. Dow is shown in fig. 3. These were verified by Dr. Houstoun, and found to agree fairly well with the values he obtained. It will be observed that generally more than two, in one case as many as four, lines were obtained to determine the point. The closeness of intersection of the different lines gives an idea of the accuracy of the method. From the directions of the lines it is possible to ascertain which auxiliary colours were used on the ground-glass plate C.

Fig. 3.



Each point took as a rule one day, as the eye required a long rest before each setting. Also the positions of the points were not the same on successive days. This was noted particularly by Dr. Houstoun in the case of the point in the extreme red, 6750 A.U., when testing whether the intensity of the light made an appreciable difference in the position of the point: of two determinations

made on different days of the same week one agreed closely with the position given on the diagram, and the other was almost twice as far out from the corner of the triangle. As the observations were made with the greatest of care and under exactly the same conditions on each occasion, and similar differences had been noticed on former occasions, there seems no doubt that the colour-vision changes appreciably from day to day.

The luminosity of the comparison fields was about 100 metre-candles when working in the red end and middle of the spectrum, but it fell off to about 5 metre-candles towards the violet end. When the luminosity was diminished to one-quarter, it was found that there was no appreciable change in the positions of the points. In the diagram, red, green, and blue are measured in units of different luminosity, the relative values of the units for the normal eye being about 100, 120, and 5 respectively. Thus the point represented by the centroid of the triangle contains about 20 times as much red as blue and 24 times as much green as blue, although it would be produced by placing equal areas of the three filters in front of a lens forming an image of a piece of ground-glass illuminated by an Osram lamp. To produce the colour containing equal quantities of the three lights, we would require to place a much larger area of blue filter in front of the lens. This colour would be represented by a point near the blue corner of the triangle. On account of this difference in the size of the units, the colour changes most rapidly in the neighbourhood of the blue corner.

As represented on the diagram, the colour of the Osram lamp is 19·6 red, 30·1 green, and 50·3 blue, and the colour of the half-watt lamp 13·3 red, 27·1 green, and 59·8 blue. Daylight from a north sky, as determined by an auxiliary experiment, was at 8 red, 20 green, and 72 blue on the diagram.

It will be observed from the diagram that the wave-lengths 6750, 6240, 6080, 5890, and 5625 are in a straight line. Then the curve bends, and the remaining points are approximately on another straight line. This is in agreement with the determinations made by previous investigators which were reduced to the same units by Dr. Houstoun in the paper already cited, only previous investigators all make the bend nearer the violet end of the spectrum, and have the curve straight from the red end almost to 5128. That is, all the points from 6750 to 5190 should be on a straight line, which is in flat contradiction to the results we have obtained.

While there is no doubt that the curve continues straight for a good part beyond 5625 and while 5300 is a point difficult to fix, there is no doubt whatever about the position of 5190.

The corners of the triangle make very good matches with 6500, 5550, and 4800 respectively.

There is no doubt whatever that the extreme violet point 4300 contains some red, and lies on the red side of the line BG. It is quite impossible to reproduce B by mixing G and 4300. The difference seems trifling on the diagram, but that is only because owing to the small luminosity of the blue unit the colours change very rapidly in the neighbourhood of B. The results that the red end of the curve is straight, while the violet end curves back towards red, are in agreement with a special investigation on these points which was carried out by Lord Rayleigh *.

It was stated in the paper already cited that the different attempts hitherto made to resolve the colours of the spectrum into primary colours are not in good agreement. As a result of our investigation, we are of opinion that this is because the subject does not admit of very exact measurement ; besides the differences which occur between different observers, the measurements made by the same observer vary somewhat from time to time. The descriptions of the previous investigations give an impression of fictitious accuracy. Newton† himself stated that the diagram was accurate enough for practice, though not mathematically accurate.

Summary

(i.) A new method of great simplicity and elegance has been worked out for evaluating the spectral colours in terms of the primary colours.

(ii.) A set of values has been obtained by this method.

(iii.) In descriptions of previous work on the subject sufficient attention has not been called to the variation in the performance of the same eye from day to day.

September 1, 1922.

* Scientific Papers, vol. v, p. 569.

† 'Opticks,' Bk. i. Pt. 2, Prop. VI. Prob. II.

XIII. *The Equality of Tensors.* *By* T. Y. Thomas *.

THE components of a tensor represent physical quantities, or else are such that physical quantities can be expressed in terms of them, as by differentiation for example. By equating the components of two tensors, relations between these quantities, *i. e.* physical laws, are obtained in a form which remains the same in all coordinate systems. The conditions under which the components of one tensor are equal to the components of another are therefore of fundamental importance. So far, these conditions have been furnished by Theorem I. of this paper. Theorem II., however, gives an entirely new set of conditions and promises to have important applications. It was suggested by a passage in Einstein's Princeton Lectures on the momentum and energy of moving bodies, and is applied to this problem in the present paper.

Definition of tensor.—A tensor of the nth order is a set of k^n functions of k variables which is transformed according to one of the following schemes :

$$\overline{\mathrm{A}}^{ab\ldots c} = \frac{\partial y_a}{\partial x_\alpha}\frac{\partial y_b}{\partial x_\beta}\cdots\frac{\partial y_c}{\partial x_\gamma}\mathrm{A}^{\alpha\beta\ldots\gamma},$$

$$\overline{\mathrm{A}}_{ab\ldots c} = \frac{\partial x_\alpha}{\partial y_a}\frac{\partial x_\beta}{\partial y_b}\cdots\frac{\partial x_\gamma}{\partial y_c}\mathrm{A}_{\alpha\beta\ldots\gamma},$$

$$\overline{\mathrm{A}}^{a\ldots b}{}_{c\ldots d} = \frac{\partial y_a}{\partial x_\alpha}\cdots\frac{\partial y_b}{\partial x_\beta}\frac{\partial x_\gamma}{\partial y_c}\cdots\frac{\partial x_\delta}{\partial y_d}\mathrm{A}^{\alpha\ldots\beta}{}_{\gamma\ldots\delta},$$

where the indices which appear twice are to be summed over the values $(1, 2, \ldots k)$. Each A carries n indices. The tensor is contra-variant, co-variant, or mixed respectively according as its scheme of transformation is the first, second, or third of these equations. The functions A are the components of the tensor in the coordinate system considered.

Definition of equality.—Two tensors of the same kind (contra-variant, co-variant, or mixed) and of the same order will be said to be equal if their corresponding components are equal in all coordinate systems.

Theorem I. If the components of a tensor A are equal to the corresponding components of a tensor B, of the same kind and order as A, in one coordinate system, then A and B are equal.

* Communicated by Prof. Oswald Veblen.

Proof.—Taking A and B to be mixed tensors as the most general case, we have

$$(\overline{A}^{a\ldots b}{}_{c\ldots d} - \overline{B}^{a\ldots b}{}_{c\ldots d}) = \frac{\partial y_a}{\partial x_\alpha}\cdots\frac{\partial y_b}{\partial x_\beta}\frac{\partial x_\gamma}{\partial y_c}\cdots\frac{\partial x_\delta}{\partial y_d}(A^{\alpha\ldots\beta}{}_{\gamma\ldots\delta} - B^{\alpha\ldots\beta}{}_{\gamma\ldots\delta}) = 0,$$

since from the conditions of the theorem,

$$A^{\alpha\ldots\beta}{}_{\gamma\ldots\delta} = B^{\alpha\ldots\beta}{}_{\gamma\ldots\delta}.$$

Corollary.—If the components of a tensor vanish in one coordinate system, they vanish in every coordinate system.

Definition of the class of a component.—The components $A^{a_1\ldots a_m a_{m+1}\ldots a_n}$ and $A^{b_1\ldots b_m b_{m+1}\ldots b_n}$ of the tensor A will be said to be in the same class if when $a_i = a_j$, $b_i = b_j$, and when $a_i \neq a_j$, $b_i \neq b_j$.

The general tensor of the nth order contains $2^n - n$ classes.

In the following we shall consider only those terms for which the number of distinct indices is not greater than k, where k is the order of the space. This restriction does not exclude any tensors of practical importance, for the components of a tensor are the coefficients of the invariant $A^{a\ldots b}{}_{c\ldots d}\xi_a\ldots\xi_b\,\xi^c\ldots\xi^d$, and the terms which involve more than k distinct indices are necessarily zero.

Theorem II. If one component α of a tensor A is equal to the corresponding component β of the tensor B, of the same kind and order as A, in all coordinate systems, then the corresponding components of the class of α and β are equal.

Proof.—From the conditions of the theorem we have

$$\overline{A}^{a\ldots b}{}_{c\ldots d} = \overline{B}^{a\ldots b}{}_{c\ldots d},$$

where $a\ldots b$, $c\ldots d = 1, 2\ldots k$ have particular values. The barred components in the coordinate system y are obtained from the unbarred in the coordinate system x by an arbitrary transformation. The unbarred components are referred to an arbitrary coordinate system, so that any relation between them will hold in all coordinate systems.

Let $\bar{\alpha}\ldots\bar{\beta}$, $\bar{\gamma}\ldots\bar{\delta}$ denote particular values of $\alpha\ldots\beta$, $\gamma\ldots\delta$ such that $A^{\bar{\alpha}\ldots\bar{\beta}}{}_{\bar{\gamma}\ldots\bar{\delta}}$ is in the same class as $A^{a\ldots b}{}_{c\ldots d}$. Then we shall have

$$A^{\bar{\alpha}\ldots\bar{\beta}}{}_{\bar{\gamma}\ldots\bar{\delta}} = B^{\bar{\alpha}\ldots\bar{\beta}}{}_{\bar{\gamma}\ldots\bar{\delta}},$$

provided that we can make a transformation such that the coefficient of

$$(A^{\bar{a}\ldots\bar{\beta}}{}_{\bar{\gamma}\ldots\bar{\delta}} - B^{\bar{a}\ldots\bar{\beta}}{}_{\bar{\gamma}\ldots\bar{\delta}})$$

is not equal to zero, while all the other coefficients are equal to zero. A transformation of this sort is

$$y_a = x_{\bar{a}}$$
$$\ldots\ldots\ \ldots\ldots$$
$$y_b = x_{\bar{\beta}}$$
$$y_c = x_{\bar{\gamma}}$$
$$\ldots\ldots\ldots\ldots$$
$$y_d = x_{\bar{\delta}}$$
$$y_e = y_e(x_1 \ldots x_k) \qquad e \neq a\ldots b,\ c\ldots d.$$

Corollary.—If one component of each class of a tensor A is equal to the corresponding component of a tensor B, of the same kind and order as A, in all coordinate systems, then A and B are equal.

Corollary 2.—If one component of a vector (tensor of the first order) is equal to the corresponding component of another vector of the same kind (contra-variant or co-variant), in all coordinate systems, then the two vectors are equal.

This follows from the fact that the components of a vector are all in the same class.

Application.—For the purpose of illustrating the above theorems, I shall include an application to the mechanics of the special theory of relativity.

Let m_0 denote the stationary mass of a particle, and x_1, x_2, x_3 its three space coordinates. Let x_4 denote the ordinary time t, and s the proper time, where we have

$$ds = dt\sqrt{1-v^2},$$

v being the velocity of the particle referred to the velocity of light as unity. The quantities $m_0 \dfrac{dx_\nu}{ds}$ $(\nu = 1, 2, 3, 4)$ have a tensor character ; written in full, they are

$$\frac{m_0 v_1}{\sqrt{1-v^2}},\quad \frac{m_0 v_2}{\sqrt{1-v^2}},\quad \frac{m_0 v_3}{\sqrt{1-v^2}},\quad \frac{m_0}{\sqrt{1-v^2}}.$$

It can be shown that the quantities I_1, I_2, I_3, E, formed from the three components of the momentum I (space components) and the energy E (time component), also possess

a tensor character *. Let us therefore make the assumption that

$$E = \frac{m_0}{\sqrt{1-v^2}}.$$

This is justified by the fact that the expansion gives $E = m_0 + \frac{1}{2}m_0v^2 + \frac{3}{8}m_0v^4 + \ldots$, in which m_0 is equal to the stationary energy E_0, and $\frac{1}{2}m_0v^2$ is the ordinary expression for the kinetic energy. To these are added terms of the fourth and higher powers of v, which are very small in comparison with $\frac{1}{2}m_0v^2$ for bodies moving with ordinary velocities. We therefore have two tensors of the first order with their fourth components equal in all coordinate systems, and hence from Theorem II. (corollary 2)

$$I_1 = \frac{m_0v_1}{\sqrt{1-v^2}}$$

$$\ldots\ldots\ldots\ldots\ldots$$

$$I_3 = \frac{m_0v_3}{\sqrt{1-v^2}}.$$

Since the force Q acting on the particle is equal to the rate of change of the momentum I, we obtain the vector equation

$$Q = \frac{dI}{dt},$$

or

$$Q_\nu = \frac{d}{dt}\left(\frac{m_0v_\nu}{\sqrt{1-v^2}}\right) \qquad \nu = 1, 2, 3.$$

Let Q_4 = the work done on the particle in unit time, then

$$Q_4 = \frac{d}{dt}\left(\frac{m_0}{\sqrt{1-v^2}}\right).$$

These are the equations of motion of the mass point.

The equations of motion for the case of a continuous distribution of matter may be obtained as follows. Let P_1, P_2, P_3 denote the three components of force on unit volume, and P_4 the work performed on unit volume in unit time. It can then be shown that P_ν ($\nu = 1, 2, 3, 4$) has a tensor character †. We have

$$P_\nu dV = Q_\nu$$

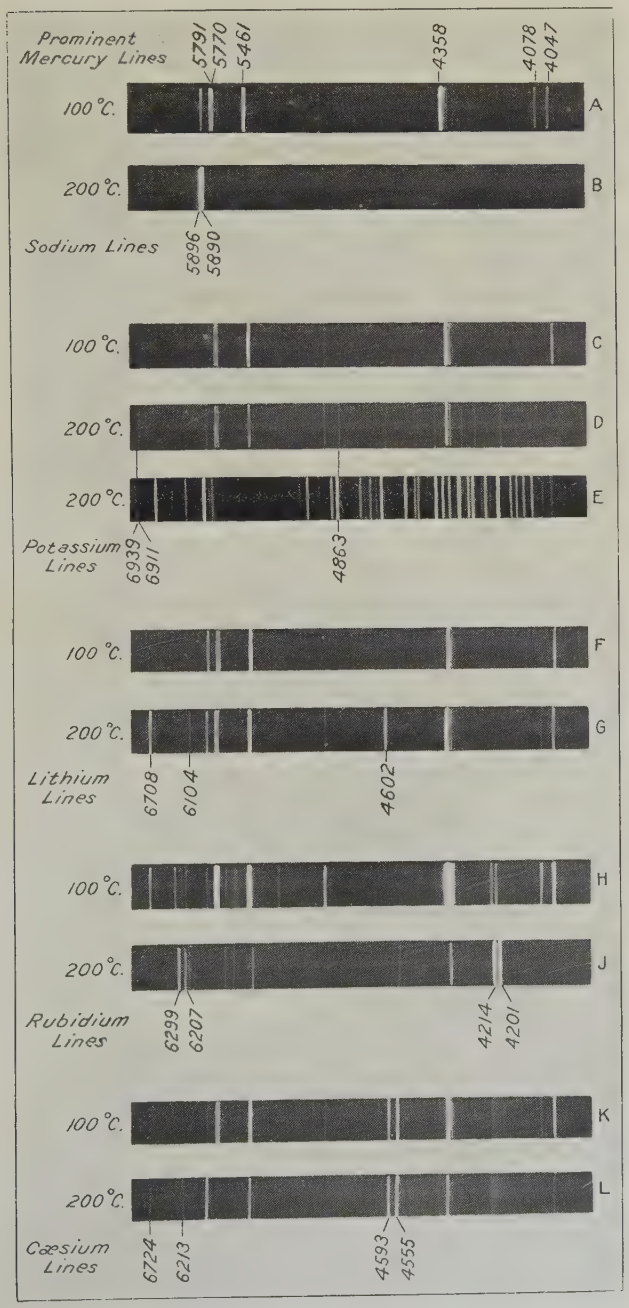
and

$$dV = dV_0\sqrt{1-v^2},$$

* Einstein Princeton Lectures.

† Weyl, *Raum-Zeit-Materie*, 4th ed. p. 150.

Spectra of Alkali Amalgams.



dV and dV_0 referring respectively to the moving and stationary volume element. From these and the four equations above for the motion of the mass point, we can obtain

$$P_\nu = \rho_0 \frac{du_\nu}{ds} \qquad \nu = 1, 2, 3, 4,$$

where

$$u_\nu = \frac{dx_\nu}{ds},$$

and ρ_0 is the stationary density of the matter. Since P_ν and $\rho_0 \frac{du_\nu}{ds}$ have a tensor character, these equations retain their form under all transformations by Theorem I.

My thanks are due to Professor Veblen, Dr. McDuffee, and Dr. Franklin for criticism and suggestions.

Princeton University,
May 1922.

XIV. *Relative Visibility of Spectra when an Electric Discharge is passed through the Vapours of Alkali Amalgams.* *By* F. H. Newman, *D.Sc., F.Inst.P., Head of the Physics Department, University College, Exeter* *.

[Plate I.]

1. Introduction.

When an electric discharge is passed through a mixture of gases or vapours, the relative visibility of the spectra of the components varies according to the conditions of the experiment. The early work of Ramsay and Collie † indicated that in a mixture of helium and hydrogen, the less helium there was present in the mixture, the lower was the pressure necessary to make its spectrum visible. The partial pressures of the components seemed to be important, and decided which of the two spectra was the brighter at a particular pressure of the whole. It is now known from Bohr's theory of the atom, that the ionization potentials of the components must be taken into account.

* Communicated by the Author.
† Proc. Roy. Soc. lix. (1896).

Atoms are capable of losing electrons when they collide with electrons possessing a definite ionizing potential, characteristic of the atom. When the electron returns to its normal position in the atom, the energy appears, in nearly all cases, as the principal series of the spectrum. Considering Bohr's assumption that radiation occurs when an electron falls from an outer to an inner orbit, combined with the assumption that complete ionization is produced much less frequently than partial ionization, *i. e.* the displacement of an electron from an orbit to one further out, it follows that radiation may occur before ionization. This resonance potential is determined by the frequency of a prominent line in the spectrum, the first line in a principal or combination series. Ionization may be brought about in ways other than by a single impact of an electron against an atom. It may arise from electron impacts against atoms in such rapid succession that the energies of the impacts are additive in effect. It may be due to a secondary effect involving combined action of radiant energy received from the surrounding atoms plus energy received by direct impact of electrons. At high pressures it is obvious that no appreciable ionization can occur as a result of single impacts, since the electrons, whenever emitted, would collide so frequently when gaining the necessary energy for ionization, that they would certainly collide inelastically and produce radiation before acquiring the energy necessary for ionization. Thus, practically all ionization observed at higher pressures, whatever be the applied potential difference, must be due to this secondary effect.

As a rule, the higher the ionization potential of an element, the greater is the difficulty with which it can be excited to emit its line spectrum. It is well known that it is difficult to excite helium ; the smallest trace of foreign gas tending to quench the helium lines. This is due to the fact that helium has the highest ionization and radiation potential of all elements, so that when it is subjected to a stimulus, this passes through the more easily excited impurities. Under most physical conditions a very small concentration of radiant atoms suffices for the production of the principal series, especially the fundamental line.

On the basis of Bohr's theory, one would expect one line after another to appear in the spectrum series as the potential is increased from the resonance potential to that of ionization. This has actually been found with mercury by Franck *, but with other vapours this is not the case.

* *Zeitschrift für Physik*, xi. (1920).

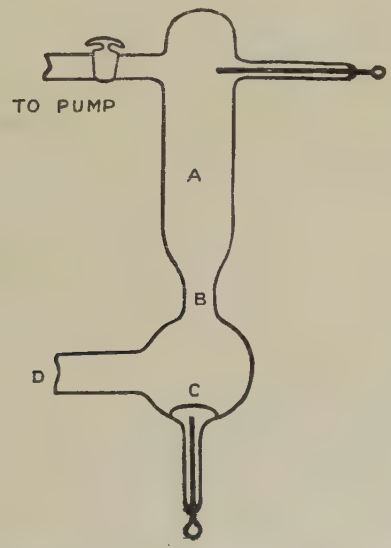
It would thus appear from the laws of energy exchange between electrons and vapour molecules that in the case of a mixture of vapours of small electron affinity, the vapour having the smaller ionization potential must always emit its spectrum to a greater degree than the other vapours.

The object of the present work was to test this theory in the cases of mixtures of the vapours of mercury and the alkali metals.

2. Apparatus.

The form of the electric discharge tube used is shown in fig. 1. It was of silica and had a constriction at B. The

Fig. 1.



mercury and the alkali were introduced, and the tube heated so that the mercury boiled. The gases liberated during the melting of the alkali were then removed by the pump. The amalgam was made the cathode, and the spectra could be viewed and photographed through the quartz window at D. In this way the light emitted at the cathode was observed. Observations also were made of the radiation emitted at B and A. The tube was enclosed in an electric heater. In obtaining the electric discharge a 10-inch induction-coil was used. It was operated with a Wehnelt interrupter joined to the 110 volt mains. Except where otherwise stated, the

amount of alkali used was always sufficient to fill the tube with its vapour at all temperatures employed. The spectrograms were photographed, by means of a constant deviation type of spectrometer, on Wratten Panchromatic plates. Some of the results obtained are shown in Plate 1. The exposures were of 60 seconds' duration except (B), where the exposure was one second only. All the spectrograms were photographed at the constriction B.

3. Mercury and Sodium.

With a mixture of sodium and mercury, the spectrum of the former was prominent at all temperatures. Below 100° C. the mercury spectrum was the brighter, but above this temperature the sodium D lines were more intense than the mercury lines, until at 200° C. the latter were almost entirely masked. A very brilliant yellow radiation was emitted at this temperature, and this method of obtaining the sodium D lines of great intensity is very convenient. It gave equally good results when the amalgam was heated with a gas-ring burner. If sodium alone is used in an electric discharge-tube, the vapour condenses on the colder parts of the tube and cuts off the radiation. No such trouble was experienced in the present work, and there was no blackening of the silica due to the action of the sodium vapour. The proximity of the mercury yellow lines 5769·6 and 5790·7 is not detrimental to the use of such an arrangement as a source of D-line radiation, as they are extremely faint, being almost entirely suppressed. The radiation is much brighter than that obtained with the form of sodium-vapour lamp previously described by the author *. The proportions of sodium and mercury used were about 1 and 100 respectively. The sodium radiation filled the discharge-tube, but was brighest at the constriction.

The amalgam was then removed and the tube thoroughly cleaned. Mercury alone was now placed in it, and on passing the discharge, at a temperature of 200° C., the constriction B was filled with the D-line radiation, the sodium, having diffused into the walls of the vessel during the previous experiment, was now re-liberated by the heat. Above and below the constriction B the radiation consisted almost entirely of the mercury lines. The amount of sodium vapour necessary for the predominance of the D-lines is very

* Proc. Phys. Soc. xxxiii. pt. 2 (1921).

minute. The experiment also shows that the current density is an important factor in determining the relative visibility of the spectra of the components of vapour mixtures.

Sodium vapour has a lower ionization potential than mercury, the values being 5·13 volts and 10·2 volts respectively. It is significant, however, that Foote and Hohler * have found that the D-lines can be emitted with a potential less than 3·3 volts, while Wood and Okano † noted that the D-lines appeared when a potential of 0·5 volt only was applied.

At low temperatures, although the vapour-pressure of mercury is large compared with that of sodium, the actual values in both cases are small. As a result, the sodium atoms in the discharge-tube will be comparatively few. The mercury atoms also will probably be so far apart that the electrons attain the energy necessary to ionize the mercury atoms before they collide inelastically with any of the atoms. At low vapour-pressures therefore, it is probable that the same proportion of mercury and sodium atoms present in tube will be ionized. The intensity of the radiation emitted must be intimately connected with the concentration of the radiating atoms in the source, so that as the concentration of the radiating mercury atoms is greater than that of sodium, we should expect the mercury lines to predominate as shown in Plate I. (A).

When the temperature is increased, the vapour-pressure of mercury rises rapidly, and the electrons will now collide inelastically with the sodium and mercury atoms before they have attained energy equal to 10·2 volts—the ionization-potential of mercury vapour,—although they may attain energy sufficient to ionize the sodium atoms. In this way the proportion of the latter which are ionized may be large, while the fraction of the mercury atoms ionized decreases ; so that although the mercury atoms present in the tube are far more numerous than the sodium atoms, the actual number of the former ionized may be considerably smaller than the number of ionized sodium atoms. The concentration of the latter is increased, and accordingly the D-lines predominate as shown in Plate I. (B). At higher temperatures very few of the mercury atoms are ionized owing to the high pressure, and the mercury lines become suppressed.

* Washington Acad. Sci. J. viii. (1918).
† Phil. Mag. xxxiv. (1917).

4. Potassium and Mercury.

At temperatures below 100° C. no trace of the potassium lines was observed, Pl. I. (C). The mercury lines were prominent except 4078, which in the case of sodium and mercury had appeared to be nearly as intense as 4047. As the temperature was increased to 200° C. the potassium lines appeared, while the intensity of the mercury lines decreased, Pl. I. (D). The former were very faint, however, and at no temperature were they as bright as the latter. The most prominent of the potassium lines were 6939, 6911, and 4863. For the sake of comparison, the spectrum obtained by passing an electric discharge through potassium vapour only at 200° C. is shown in Pl. I. (E).

The theory that a vapour having a low ionization potential should always emit its spectrum to a greater extent than one having a high ionization potential, apparently fails in this case, for the ionization potential of potassium vapour is 4·32 volts—lower than that of sodium. Even if the intrinsic brightness of the sodium D-lines accounts for the brilliance of the sodium radiation, it would be expected that the potassium line 4047 would be fairly prominent, but such is not the case. It has been shown by the author in the paper already quoted that in the case of a mixture of sodium and potassium vapours, the spectrum of the latter is easily suppressed.

The reason for the faintness of the potassium lines is not clear, but it is significant that Hebb * found that the mercury spectrum could be produced at a potential as low as 0·5 volt in an atmosphere of mercury and potassium. There is a decrease in the brightness of the mercury spectrum when the temperature increases, as there should be according to the theory developed in the case of sodium and mercury.

The smallest trace of mercury vapour in the tube results in the appearance of the mercury lines: and they are much brighter than those of potassium.

5. Lithium and Mercury.

Below 100° C. the lithium lines do not appear, Pl. I. (F), but at 200° C. they are very prominent, although never showing much tendency to mask the mercury lines, Pl. I. (G). The lithium lines are most intense at the cathode, and the constriction of the tube. The most prominent line is 6708, although 4602 is strong. The mercury line 4078 is suppressed.

* Phys. Rev. xii. (1918).

The ionization potential of lithium vapour calculated from the quantum theory is 5·38 volts, higher than that of sodium vapour. The vapour-pressure of lithium, at the temperatures used, is very low. These two facts account for the lithium lines not being as intense as might be expected at 200° C.

6. Rubidium and Mercury.

At temperatures below 100° C. the mercury spectrum predominates, although the rubidium doublet 4201, 4214—the second doublet of the principal series—is always present, Pl. I. (H). As the temperature is raised to 200° C. the colour of the radiation in the tube changes from the characteristic green of mercury to the lilac of rubidium. At this temperature the spectrum of the latter is very bright, the mercury lines being almost suppressed, Pl. I. (J). Above 200° C. the rubidium lines were still very intense compared with the mercury lines. Only in the cases of sodium, rubidium, and cæsium is the mercury green radiation masked by the radiations from these metals. Potassium and lithium do not show this effect.

The ionization potential of rubidium vapour is low—4·1 volts *—and its vapour-pressure is greater than that of sodium, lithium, and potassium. These two facts account for the greater intensity of the rubidium lines compared with those of lithium and potassium. The flame lines 6299, 6207 are present.

In this case there was considerable blackening of the quartz tube due to the action of rubidium vapour.

The suppression of the mercury spectrum, and increased brightness of the rubidium lines as the temperature increases, can be explained by a theory similar to that advanced in the case of sodium.

7. Cæsium and Mercury.

Below 150° C., the mercury lines are the most intense in the spectra of a mixture of mercury and cæsium vapours, although the cæsium doublet 4593, 4555 is always prominent, Pl. I. (K). The radiation in the tube shows the characteristic green colour of mercury, but above this temperature the colour changes to purple, and at 200° C. it is pink. The latter is characteristic of cæsium radiation. At this temperature, the mercury lines are less intense than the cæsium lines, Pl. I. (L). This is very marked at the constriction of the tube. Cæsium lines, other than the doublet

* Franck & Hertz, *Verh. d. Deutsch. Ges.* xx. (1919).

in the blue, make their appearance—6724, 6213,—but they are comparatively faint. The mercury line 4078 is again suppressed, Pl. I. (K).

The vapour-pressure of cæsium at any particular temperature is greater than that of the other alkali metals at the same temperature. Its ionization potential (3·9 volts) is also low. It is to be expected, therefore, from what has been said previously, that the cæsium lines will increase in brightness compared with the mercury lines, as the temperature is increased.

There was considerable "browning" of the quartz-tube due to the action of the cæsium vapour on the silica.

8. Discussion of Results.

The experimental results obtained indicate that two factors, in general, determine the relative visibility of the spectra of the components when an electric discharge is passed through a mixture of gases or vapours : (*a*) the ionization potential of the vapour, (*b*) the partial pressure of the vapour. The intensity of radiation depends on the radiant energy emitted in unit time. The amount of energy emitted during each process in which emission occurs can be calculated from the quantum relation $E = h\nu$, and this amount will be one determining factor of the intensity of the spectrum lines. Another factor will be the frequency with which the process occurs. At very low pressures in the discharge-tube the electrons collide so seldom when gaining the energy necessary to produce ionization on impact with the atom, that they do not collide inelastically before acquiring the energy necessary for ionization, providing the latter is not very large. Taking the example of sodium and mercury vapours already discussed, although there is greater probability of the sodium atoms being ionized because of their small ionization potential, at very low pressures, the electrons probably acquire energy of 10·2 volts before collision, and as a result, the chance that the requisite energy will be available for the emission of the sodium spectrum will be practically equal to the chance that it will be available for the mercury. In such circumstances we should expect that the concentration of the radiating atoms would be greater for that vapour having the greater partial pressure, *i. e.* the one having the greater number of atoms per unit volume.

When the pressure becomes greater, there is less probability of the electrons attaining large energy before collision. The maximum energy attained may be sufficient to ionize atoms of small ionization potential, such as those of sodium, but not sufficient to ionize the mercury. Although the atoms of

higher ionization potential may be more numerous, the fraction of them which radiate energy may be very small. On the other hand, there may be comparatively few atoms of the element of lower ionization potential present, but a large percentage of them will be radiating atoms. In this case we should expect the spectrum of the element of lower ionization potential to be the more prominent. The experiments described above have shown this to be true.

This theory will hold if the ionization takes place owing to single impacts by the electrons, or if the energy necessary for ionization is accumulated from successive impacts and the effect of radiation from adjacent atoms.

9. Summary.

1. The relative intensity of the spectra produced when an electric discharge is passed through a mixture of the vapours of mercury and the alkali metals has been investigated.

2. At 100° C., the mercury lines predominate, but at 200° C. the spectra of the alkali metals are brighter than the mercury spectrum. The sodium D-lines mask the mercury lines at the higher temperature. The potassium spectrum is very faint under all conditions.

3. An explanation of the effect is based on the low vapour-pressures of the alkali metals at these temperatures, and on the low values of the ionization potentials which the vapours of sodium, potassium, lithium, rubidium, and cæsium have.

4. The employment of the amalgams of the alkali metals, instead of the latter, forms a useful and convenient method for obtaining the spectra of the alkali metals. A brilliant D-line radiation is obtained with sodium amalgam.

XV. *On the Theory of the Specific Heat of Liquids.* By Kôtarô Honda*.

1. THE theory of the specific heat of gaseous and solid phases is already well known. The results deduced from the kinetic theory of gases agree satisfactorily with the observed facts, provided the temperature is not extremely low. The theory of the specific heat of solids first developed by A. Einstein †, and afterwards improved by Nernst ‡, Debye §, and others, accords with the observed facts quite satisfactorily, provided the temperature is not very high. In

* Communicated by the Author.

† Einstein, *Ann. der Phys.* xxii. p. 180 (1907).

‡ Nernst & Lindemann, *Zeitschr. für Elektrochemie*, 1911, p. 817; *Berl. Ber.* 1910, p. 26.

§ Debye, *Ann. der Phys.* xxxix. p. 789 (1912).

this case the atomic heats of many metals are a little greater than the theoretical value 5·96. As regards the liquid phase, we have as yet no satisfactory theory; the following is an attempt to fill the above gap.

Before entering upon the theory of the specific heat of the liquid phase, it is necessary to explain the mechanism of the melting of solid substances. For the sake of simplicity we take the case of the monatomic substance, which has the same atomic form both in the liquid and solid phases. Since the solid is always a crystal or crystalline, the atoms are distributed in the space-lattice belonging to its crystal system. According to the modern view, the atom has a very complicated structure, so that it sends out its lines of force chiefly in certain definite directions, and hence the regular spacing of atoms in a solid is only possible for their definite orientations—that is, in the solid substances, free rotation of the atoms is not possible. This view is confirmed by the determination of the specific heat of solid substances. Thus each atom of a solid vibrates about its mean position with three degrees of freedom.

We may also with good reason suppose that with the above mode of rectilinear vibration the atoms are associated with small dependent rotational vibrations about their centres of mass, probably with a common period of vibrations. At the ordinary temperature this mode of rotational vibration is very small; but as the temperature rises, it increases at a steady rate. The fact that at very high temperatures the specific heat experimentally found is somewhat greater than the theoretical value 5·96 indicates the presence of this rotational vibration On approaching the melting-point of the substance, the amplitude of this vibration attains such a value that the rotational vibration at last becomes a continuous rotation. As soon as the rotational vibration is changed into a continuous rotation, the regular spacing of the atoms breaking down, disorder begins to ensue, and therefore the rotation becomes independent of the translation. The atoms thus set free have now two or three degrees of freedom for rotation; the angular velocity of each component rotation increases to such an extent that its energy becomes equal to that corresponding to one degree of freedom of rotation. If all the atoms undergo the above change, the fusion is said to be completed, and the crystal changes into a liquid. The above is more satisfactory from a physical point of view than the view proposed by Lindemann* regarding the melting.

If the above view of fusion be correct, the heat of fusion

* Lindemann, *Phys. Zeitschr.* xi. p. 609 (1910).

consists mainly of the energy of rotation of the atoms gained during the process of fusion, the potential energy of atoms due to the small expansion of the volume during fusion being comparatively small.

Since in solids the motion of each atom is vibration, but not translation as in the case of gases, its total energy is just twice the mean kinetic energy during a complete period of vibration. Hence, according to the quantum theory, the energy belonging to each degree of freedom is given by

$$\epsilon_T = \frac{h\nu}{e^{\frac{h\nu}{rT}} - 1},$$

where T is the absolute temperature, r the gas constant referred to one atom, h the universal constant, and ν the frequency of vibration. During the process of fusion each atom gains an energy of rotation corresponding to two or three degrees of freedom, the energy for each freedom of rotation being $\frac{1}{2}\epsilon_{T_m}$, where T_m is the melting-point. If one gram atom contains N atoms, the total energy of rotation gained by the atom is

$$N\epsilon_{T_m} \ (=E_{T_m}) \quad \text{or} \quad \tfrac{3}{2}N\epsilon_{T_m} \ (=\tfrac{3}{2}E_{T_m}),$$

according as the freedom of the liquid atom is 2 or 3. Since the atomic distances in the solid and liquid phases do not differ much from each other, the rectilinear motion of the liquid atoms is also vibratory, its amplitude remaining unchanged. Hence the translational energy of each liquid atom for one degree of freedom will also be equal to E_{T_m}—that is, the same as that of the solid state or very nearly so. In general, the change of volume during melting is very small, so that the energy required for doing the internal and external work is also small; we denote it by Q. If λ be the latent heat of fusion of a substance and w its atomic weight, we have

$$w\lambda = \tfrac{1}{2}nE_{T_m} + Q,$$

where $n=2$ or 3. If we neglect Q in this relation, we have simply

$$w\lambda = \tfrac{1}{2}nE_{T_m} \quad \text{or} \quad \frac{w\lambda}{T_m} = \frac{nE_{T_m}}{2T_m}.$$

Generally E_{T_m}/T_m has a value a little less than $R=1{\cdot}96$, the gas constant referred to one gram atom, the calorie being the unit of heat. This simple relation has been shown by the present writer * to be satisfied in the case of all metals in which λ and ν are known.

In the above view it is assumed that in the liquid phase even the mono-atom has two or three degrees of freedom for

* K. Honda, Sci. Rep. vii. p. 123 (1918); Phys. Rev. xii. p. 425 (1918).

rotation. This assumption is, however, quite natural, because, according to the electron theory, the atom itself has a very complicated structure, being quite different in shape from a sphere; but in a gaseous phase, the atomic distance being very large as compared with its dimensions, such an atom may behave like a sphere.

If the melting-point of a substance is sufficiently high to admit the equipartition of energy, we have then

$$E_{T_m} = R_{T_m};$$

hence the total energy of the substance at a temperature T above its melting-point is

$$W = 3RT_m + \tfrac{1}{2}nRT + Q + 3R(T - T_m)(1 - e),$$

where e has the value lying between 0 and $\frac{1}{2}$; for, if each atom in the liquid phase is making a simple vibration as in the case of the solid phase, we must have $e = 0$. On the other hand, if the temperature be so high that each atom has a translational motion like a gas molecule—that is, in its almost entire length of the free path each atom has a uniform translational motion, and only in very short trips at both ends a retarded or accelerated motion,—then the energy of translational motion for three degrees of freedom is $\frac{3}{2}RT$, and not 3RT as in the case of a simple vibration; hence in this case we have $e = \frac{1}{2}$. In the actual case, each atom in the liquid phase has a rectilinear motion of a kind lying between the above two cases, and hence $0 < e < \frac{1}{2}$.

The atomic heat at constant volume in the liquid phase is then

$$C_v = \frac{dW}{dT} = 3R(1 - e) + \tfrac{1}{2}nR.$$

Hence the values of C_v lie in the following ranges :—

n =	2,	3,
C_v =	4·95—7·92,	5·94—8·91.

Since e may increase very slightly with the rise in temperature, C_v is almost independent of temperature, and in some cases slightly decreases with it.

If C_v' be the atomic heat in the solid phase at the melting-point, we have

$$C_v - C_v' = R\left(\frac{n}{2} - 3e\right),$$

which is the change of atomic heat during melting. This difference has a positive or negative value according to the nature of the substance.

The above conclusions agree with the facts experimentally found*.

* I. Iitaka, Sci. Rep. viii. p. 99 (1919).

XVI. *On Castigliano's Theorem of Least Work, and the Principle of St. Venant.* *By* R. V. Southwell *.

[From the National Physical Laboratory.]

Introduction and Summary.

1. A "FRAMEWORK" is a skeletal structure, designed to transmit through its component members a system of external forces applied at its joints. We shall assume in this paper that every member would be free to turn about either of its ends if the other end were released. Under these conditions, the action in each member must be a pure tension or compression : it is an ideal which practical design seeks generally to attain, since the material can thus be used to the best possible advantage.

A "simple framework" is one which would collapse if any single member were removed : that is to say, it has sufficient members, but no more, to maintain the relative positions of its joints. When the number of members is more than sufficient for this purpose, the structure is termed "redundant". Sometimes it will happen that a framework, although apparently redundant, includes several members which are too flexible to withstand any appreciable compression, and that enough of these go out of action, under the load-system considered, to make it, in effect, simple : the framework is then said to be "pseudo-redundant".

In a redundant framework, it is often convenient to distinguish between two sets of members—the first consisting of "essential members", sufficient by themselves to maintain the position of every joint, and the second of "redundant members". Such classification will of course be largely arbitrary.

2. Redundancy has several important consequences. From the definition of a simple framework it follows that the actions in its members can be found, in terms of the applied loads, from statical considerations alone ; in a redundant framework the actions depend also upon the elastic properties of the members. Again, if the external forces be removed from a simple framework, no stress will be left in any member ; in a redundant framework, stresses can occur independently of the external loads, by "self-straining", if the length of any redundant member, before insertion in the framework, is not exactly equal to the distance between the joints which it has to connect. Self-straining is often imposed deliberately, by means of turnbuckles, etc., which

* Communicated by the Author.

enable the lengths of redundant members to be altered after insertion. In this way (since a wire which is initially tensioned can function, until its tension is entirely annulled, as a strut), a pseudo-redundant framework can be rendered truly redundant, and its stiffness thereby increased.

From the practical standpoint, a redundant framework has the disadvantage that the stresses in its members are harder to calculate, and the results (for the reasons indicated) less reliable; but for some purposes, and especially in aircraft, its increased stiffness is a more than compensating advantage. Although in the ideal framework the joints are assumed to be free, yet in practice such freedom is often difficult to ensure, and in rigid airships, for example, the joints are more or less rigidly "fixed". Under these conditions, the greater the flexibility of the structure the greater are the racking actions imposed upon the joints, which are almost always points of weakness. For these and other reasons, the problem of calculating stresses in redundant structures is increasing in practical importance.

3. The obvious and direct method of attack is to "compare displacements" *. Considering the simple framework formed by the "essential" members, we can calculate its deflexion under the given load-system, and the additional displacements which result when a specified force is imposed upon it by any one of the redundant members. Hence, regarding the actions in the redundant members as unknowns, and making use of the condition that the extension of any redundant member must be equal to the relative displacement of the joints which it connects, we can obtain sufficient equations for a solution.

This method is (theoretically) applicable, even though the load-extension relations for the members are not expressed by Hooke's law. It has the disadvantage, however, that the calculation of displacements (at all events in a three-dimensional framework) is a lengthy and troublesome operation. A systematic process for effecting this, in structures which satisfy Hooke's law, was described by Clerk Maxwell, in a paper communicated to this Magazine in 1864 †; but the method, in spite of

* The practical application of this method has been fully described by W. S. Farren, "The Calculation of Stresses in a Redundant Structure by the Method of Comparison of Deflexions, etc.", Aeronautical Research Committee, R. and M. 769.

† "On the Calculation of the Equilibrium and Stiffness of Frames", Phil. Mag. vol. xxvii. pp. 294–299. Owing, probably, to the extremely condensed exposition employed, this paper attracted little notice until the method had been rediscovered by Mohr in 1885. *Cf.* the preface to Mr. Farren's paper (*loc. cit.*), contributed by Prof. H. Lamb.

its formal elegance, involves very serious labour in practice, and redundant frameworks having extension in three dimensions are more usually treated by a method which depends upon Castigliano's celebrated "Theorem of Least Work". For present purposes, this may be stated as follows:—*In a framework which is not initially self-strained, the stresses imposed by a given system of external forces may be found from the conditions for a minimum value of the total strain-energy, if account be taken of the conditions of equilibrium.* It is restricted to structures which obey Hooke's law.

The great merit of this theorem, from a practical standpoint, is that it enables us to dispense with any detailed consideration of displacements. In the equations of equilibrium we have, of course, to deal with forces and actions as vector quantities; but this is a comparatively simple matter, and the subsequent calculations are concerned only with energy,—*i.e.*, with scalar quantities.

4. Castigliano's first statement of his theorem appeared in a thesis for the diploma of engineering at Turin, in 1873. The proof consisted simply in a direct comparison of calculated results.

In a subsequent memoir * he showed that the theorem is a direct corollary of another (also contained in his thesis of 1873), which may be stated thus:—*If the total strain-energy stored in a framework be expressed as a function of the external forces, then the partial differential coefficient of this function with respect to any one of the forces gives the relative displacement of the joint at which that force is applied* †. If the action in any one redundant member be regarded as an external force applied to the remainder of the framework, this theorem gives the expression

$$\delta_1 = \frac{\partial V_1}{\partial P} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

for the relative approach of the joints which are connected by the member considered: P denotes the tension in the member, and V_1 the total strain-energy stored in all the remaining members. In the same way, if we regard

* "Nuova teoria intorno all' equilibrio dei sistemi elastici," Trans. Acad. Sci. Turin, vol. x. (1875), pp. 380–423.

† A proof of this theorem, generally known as the "First Theorem of Castigliano", is given below (§ 12).

the action P as an external force applied to the member in question, and denote the strain-energy stored in this member by V_2, we obtain the expression

$$\delta_2 = \frac{\partial V_2}{\partial P} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

for its extension under load. Hence, if the member, before insertion in the unstrained framework, was short by an amount λ of the length required to enable it to fit into place, then the condition which must be satisfied, when the member is fitted and the forces applied, is

$$\begin{aligned} \lambda &= \delta_1 + \delta_2, \\ &= \frac{\partial}{\partial P}(V_1 + V_2), \\ &= \frac{\partial V}{\partial P}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3) \end{aligned}$$

where V denotes the strain-energy stored in the complete framework. If the framework was unstrained before the given load-system was applied, λ is zero in this equation, and it may be interpreted as one of the conditions for a minimum value of the total strain-energy V. Corresponding to every unknown tension P we have one equation of the type (3), and a solution can therefore be obtained.

This argument was extended in a treatise ('Theorem de l'Equilibre des Systèmes Elastiques et ses Applications') published in 1879*. By writing equation (3) in the equivalent form

$$\frac{\partial}{\partial P}\{V - \Sigma[\lambda P]\} = 0, \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where the summation extends to every redundant member†, Castigliano showed that the stresses in a redundant framework, of which the self-straining is defined by specified quantities of the type λ, may be determined from the

* This treatise has recently been translated by E. S. Andrews ('Elastic Stresses in Structures', 1919). Since the original edition is somewhat scarce, all references in this paper have been made to the English edition: the foregoing historical notes have been taken from Castigliano's preface and from Chap. I. of the translation.

† It can equally be taken as extending to *every* member, since λ is zero for each "essential member".

condition for a minimum value of the quantity

$$V-\Sigma[\lambda P], \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

if account be taken of the conditions of equilibrium. This theorem obviously embraces the Theorem of Least Work.

He showed, further, that analogous theorems hold in respect of the stresses in a continuous elastic solid *. Similar generalization is permissible in regard to the arguments of this paper: it is only for the sake of simplicity that the detailed calculations have been confined to frameworks with free joints.

5. While the advantages of the "Method of Least Work", in problems relating to redundant frameworks, are now generally recognized, there is a widespread feeling that the theorem underlying it is distinctly artificial, and its principles hard to understand †. An alternative proof of the theorem, on physical rather than mathematical lines, may therefore be of interest. In Castigliano's proof, the idea of a minimum appears merely in the mathematical interpretation of equations obtained by comparing displacements, and the device by which equation (3) is thrown into the form (4) is even more artificial. In the argument now advanced, the tendency of the stored strain-energy towards a minimum appears rather as an essential property of materials which obey Hooke's law. Further, if the existence of this property be admitted, it is shown that the well-known "Principle of St. Venant" follows as an immediate corollary.

The argument falls into two parts. In the first, Castigliano's theorems are shown by general reasoning to follow directly from the two principles (i.) of superposition and (ii.) of conservation of energy: at the suggestion of Prof. H. Lamb, a simple illustrative example has been added at each stage of the argument. In the second, these conclusions are confirmed by exact analysis.

The author desires to acknowledge his indebtedness to Prof. B. M. Jones, at whose suggestion this paper was written, and to Prof. H. Lamb, to whom, as on many former occasions, he is indebted for kindly criticism and valuable advice.

* "A general proof is in fact at once obtained if we regard the symbols in his equations as standing for generalized coordinates instead of extensions of bars" (H. Lamb, *loc. cit.*).

† *Cf.* W. S. Farren, *loc. cit.* p. 7, and the preface by Prof. H. Lamb.

Part I.

An Alternative Proof of the Theorem of Least Work.

6. It should be emphasized at the outset that Castigliano's theorems depend upon the validity of Hooke's law, as an expression not only of the stress-strain relations for each element of the material, but also of the load-extension relation for each member as a whole *. If this requirement is satisfied, the principle of superposition is an immediate consequence: that is to say, any given load-system applied to a framework, and in equilibrium, will increase the actions in the different members and the relative displacements of the joints by amounts which are independent of the stresses already existent. It is immaterial whether these stresses are due to external forces or to self-straining †.

Hence, if all the forces of the system are increased in any constant proportion, the corresponding internal stresses and the displacements of the joints will increase in the same ratio. Again, if these increases occur in a slow and continuous manner, so that equilibrium is established at every stage, and if the system comprises all the external forces which act upon the framework ‡, the work done by each force will be measured by the product of its final value and of half the final displacement, along its line of action, of the joint at which it acts. The total work done on the framework will thus be unaffected by any stresses which may initially exist in it as a consequence of self-straining.

By the principle of conservation of energy, the work thus done by the applied forces must all be stored as strain-energy in the framework. If, then, we write

$$V = V_0 + V', \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

where V_0 denotes the total strain-energy stored in a framework *before* the given load-system is applied, and V the total strain-energy *after* the system is applied, our argument

* Caution is therefore necessary in dealing with structures in which the same member acts both as a compression member of the framework and as a girder transmitting laterally-applied loads to the joints: in such a member, the relation between end-load and total extension may be only approximately expressed by Hooke's law, even when this law represents the relation between the stress and strain in every component element. *Cf.* A. J. S. Pippard, "The Reduction of the Effective Value of Young's Modulus in Flexible Compression Members", Aeronautical Research Committee, R. and M. 792.

† *Cf.* § 10.

‡ If other forces act initially, they also will do work as the loads of the given system are increased, and the argument will be modified accordingly.

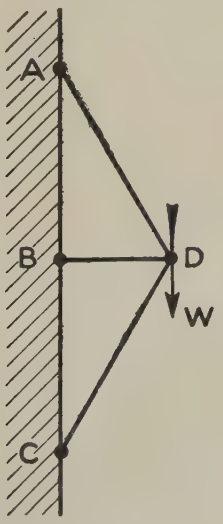
shows that V′, the total increase due to the applied load-system, depends only on that system and on the elastic properties of the framework: it is independent of V_0*.

Now V_0, V′, and V, in the nature of the case, are essentially positive quantities; hence, if V′ be given, the value of V is least when V_0 is zero. In other words, *a framework which is initially unstrained will have less strain-energy stored in it when subjected to any definite load-system than it would have if self-strained to begin with.*

But this, in effect, is the Theorem of Least Work. For an expression obtained by means of the conditions of equilibrium, for the total strain-energy in terms of the actions in the redundant members, will obviously embrace every stress-system which can occur in the framework under given applied forces, when no restriction is imposed upon the nature and amount of the initial self-straining; and the result just established shows that the conditions for a minimum value of this expression will yield the correct solution for a framework which is initially unstrained.

A simple example will serve to illustrate the identity stated in (6). In fig. 1, the members DA, DB, DC are

Fig. 1.



connected by smooth joints to the rigid wall ABC: they are assumed to be of uniform material and cross-section, such

* *Cf.* § 13.

that the member BD, which is 100 ins. long, would extend by $\frac{1}{10}$ in. under a tension of 1000 pounds weight, and the members AD and CD, which are 200 ins. long, by twice that amount. It is easily verified that the self-straining involved when BD is initially short by $\frac{1}{20}$ in. consists of 100 pounds tension in BD, together with compressions of the same amount in AD and CD; also, that a load of W pounds applied at D, as shown, will introduce a tension of $W/\sqrt{3}$ pounds in AD and an equal compression in CD, leaving BD unstrained.

Then if BD is initially short by the unknown amount x ins., the energy stored by self-straining only is as follows:—

in BD,

$$\tfrac{1}{2}\times 2000x\times\frac{2x}{10} \qquad = 200x^2 \text{ in.-pounds;}$$

in AD or CD,

$$\tfrac{1}{2}\times(-2000x)\times\left(-\frac{4x}{10}\right): \text{Sum} = 800x^2 \text{ in.-pounds.}$$

$$\text{Total} \qquad = 1000x^2 \text{ in.-pounds} = V_0.$$

When the load W is added, the strain-energy is altered as follows:—

in BD,

$$\tfrac{1}{2}\times 2000x\times\frac{2x}{10} \qquad = 200x^2 \text{ in.-pounds, as before;}$$

in AD,

$$\tfrac{1}{2}\times\left(\frac{W}{\sqrt{3}}-2000x\right)^2\times\frac{2}{10^4},$$

in CD,

$$\tfrac{1}{2}\times\left(-\frac{W}{\sqrt{3}}-2000x\right)^2\times\frac{2}{10^4}:$$

$$\text{Sum} = \frac{2W^2}{3\times 10^4}+800x^2 \text{ in.-pounds.}$$

$$\text{Total} \qquad = \frac{2W^2}{3\times 10^4}+1000x^2 \text{ in.-pounds,}$$
$$= V.$$

Thus the increase $V' = V - V_0 = \frac{2W^2}{3\times 10^4}$ in.-pounds,—a quantity independent of x (or of V_0).

Extension to Frameworks which are Initially Self-Strained.

7. Similar reasoning can be employed to establish Castigliano's theorem in its general form. By the principle of superposition, if the redundant members of a framework be lengthened by definite amounts (*e.g.*, by relaxing the turnbuckles mentioned in § 2), the tension in every member will be reduced by an amount which does not depend either upon the initial self-straining or upon the external forces. But if the framework was initially unloaded, and self-strained to an extent which is exactly neutralized when the redundant members are lengthened by the specified amounts, the tension in every member will be reduced to zero. Hence, in the self-strained framework, if the redundant members are lengthened by amounts λ_1, λ_2, ... etc., the tensions in the members will be reduced by amounts P_0, ..., etc., where P_0, ... etc. are the tensions induced by self-straining when the framework is unloaded and the redundant members "short" (as described in § 4) by the amounts $\lambda_1, \lambda_2, \ldots$ etc.

Let us now imagine that we are given the forces applied to a framework, and the amounts, λ_1, λ_2, .. etc., by which the redundant members are initially "short". Using the conditions of equi ibrium, we can calculate the action in every member in terms of the applied forces and of the actions in the redundant members, and a corresponding expression for the total strain-energy V is at once obtainable; for if we denote the tension in any member by P, and define its elastic properties by means of the equation

$$P = \Omega l^3 e \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

(where Ω is a constant for the member *, l its length, and e its extension under the tension P), we have

$$V = \Sigma[\tfrac{1}{2}Pel] = \tfrac{1}{2}\Sigma\left[\frac{P^2}{l^2\Omega}\right], \quad . \quad . \quad . \quad . \quad (8)$$

the summation extending to every member of the framework.

If we now lengthen the redundant members by amounts equal to our specified values of $\lambda_1, \lambda_2, \ldots$ etc., the tension in any member will be reduced, as shown above, by a definite amount P_0. After this operation, the total strain-energy will therefore be given by

$$V'' = \tfrac{1}{2}\Sigma\left[\frac{(P-P_0)^2}{l^2\Omega}\right]. \quad . \quad . \quad . \quad . \quad . \quad (9)$$

* *Cf.* equation (12), § 9, and footnote.

As in § 6, the expression obtained for V, since it contains the actions in the redundant members as unknowns, will embrace every stress-system which can occur in the framework, under the given forces, when no restriction is imposed upon the nature and amount of the initial self-straining. The expression (9) is equally general. But if the correct values are inserted for the actions in the redundant members, so that (8) corresponds to the correct system of self-straining, then after the redundant members have been lengthened the framework will be entirely free from self-straining, and hence, by the Theorem of Least Work, the total strain-energy stored in it when subjected to the given load-system will have its least possible value. We can therefore derive the correct solution from the conditions for a minimum value of V″.

The form of the expression (9) is not a convenient one for this calculation, since we should require to know P_0, ... etc. in terms of λ_1, λ_2, ... etc. We therefore investigate an alternative expression for the quantity $(V-V'')$. By the principle of conservation of energy, this quantity must be equal to the work gained in the operation of increasing the lengths of the redundant members, less the work done during this operation by the external forces; for the work stored by the external forces (and recoverable by removing them) will be unchanged. The tension in any one of the redundant members is represented by P at the start and by $(P-P_0)$ at the finish, so that its mean value during the operation is $(P-\frac{1}{2}P_0)$: hence, the work gained in lengthening this member is $(P-\frac{1}{2}P_0)\lambda$. Again, if u, v, w, ... are the displacements of the joints from the zero configuration * when the framework is self-strained to an extent defined by λ_1, λ_2, ... etc., and if $\overline{X}$, $\overline{Y}$, $\overline{Z}$, ... etc. are the external forces acting on the framework, the total work done by these forces in the operation (which *reduces* the displacements by amounts u, v, w, ...) will be given by the expression

$$-\Sigma\left[u\overline{X}+v\overline{Y}+w\overline{Z}\right],$$

in which the summation extends to every joint.

We thus obtain the expression

$$V-V''=\Sigma_m\left[(P-\tfrac{1}{2}P_0)\lambda\right]+\Sigma_j\left[u\overline{X}+v\overline{Y}+w\overline{Z}\right], \quad (10)$$

in which Σ_m denotes a summation extending to every member, and Σ_j a summation extending to every joint.

* *I. e.*, from their positions in the simple and unstrained framework formed by the "essential members" only. *Cf.* § 9.

But the terms on the right of (10), excepting those of type $P\lambda$, are invariant, since they depend only upon the external forces and upon λ_1, λ_2, ... etc. Hence, the conditions for a minimum value of V'' are the conditions for a minimum value of the quantity

$$V-\Sigma[P.\lambda]; \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

and since this is the quantity similarly denoted in equation (5), we have obtained an alternative proof of Castigliano's theorem in its application to self-strained frameworks.

The relation expressed by (10) may be illustrated by means of the example previously considered. Suppose, as before, that BD is initially short by the unknown amount x inches, and that a load of W pounds weight is applied at D; and let us imagine that BD is now relaxed (*i. e.*, lengthened) by $\frac{1}{20}$ inch. The effect is to superpose a compression of 100 pounds in BD and tensions of equal magnitude in AD and CD, whence the strain-energy stored after the operation is as follows :—

in BD,

$200(x-\frac{1}{20})^2 = 200x^2-20x+\frac{1}{2}$ in.-pounds;

in AD,

$\frac{1}{10^4}\times\left(\frac{W}{\sqrt{3}}-2000x+100\right)^2$,

in CD,

$\frac{1}{10^4}\times\left(-\frac{W}{\sqrt{3}}-2000x+100\right)^2$:

$$\text{Sum}=\frac{2W^2}{3\times 10^4}+800x^2-80x+2 \text{ in.-pounds.}$$

$$\text{Total} = \frac{2W^2}{3\times 10^4}+1000x^2-100x+\frac{5}{2} \text{ in.-pounds,}$$
$$= V''.$$

It follows that

$$V-V'' = 100x-\tfrac{5}{2} \text{ in.-pounds.}$$

But the tension in BD was $2000x$ pounds at the start and $(2000x-100)$ pounds at the finish, so that $P=2000x$ and $P_0=100$; also $\lambda=\frac{1}{20}$. Hence we have

$$(P-\tfrac{1}{2}P_0)\lambda=(2000x-50)\times\tfrac{1}{20}=100x-\tfrac{5}{2}=V-V'',$$

and the relation (10) obtains, since in this example no work

is done by the force W during the process of lengthening BD.

The effect of lengthening AD or CD may be investigated in the same way : it will be found that (10) is still satisfied, although work is here done by W.

A Physical Basis for the Principle of St. Venant.

8. If in the foregoing proof of Castigliano's theorem we regard the symbols as standing for generalized coordinates *, it appears that the stresses in any elastic solid which obeys Hooke's law will adjust themselves, under a given system of applied tractions, so that the total strain-energy stored by those tractions has the least value consistent with the satisfaction of the conditions of equilibrium.

If we regard this tendency as in the nature of a physical property of the material, we may find in it the basis of the principle, due to St. Venant, which is known as the "Elastic Equivalence of Statically Equipollent Systems of Load". According to this principle †, which is widely utilized in the theory of elasticity, the strains which are produced in a body by the application, to a small part of its surface, of a system of forces statically equivalent to zero force and zero couple are of negligible magnitude at distances which are large compared with the linear dimensions of the part ; hence, by the principle of superposition, it may be inferred that *forces applied at one part of an elastic structure will induce stresses which, except in a region close to that part, will depend almost entirely upon their resultant action, and very little upon their distribution.*

St. Venant's principle appears to have been originally propounded, and to have remained since, without proof. But the theorem just stated in italics may be regarded as an immediate consequence of the tendency towards a minimum storage of strain-energy. Let us consider the case of a long girder, or braced framework, which is loaded at its two ends by forces applied in any given way ; and let us employ the symbol A to denote those regions which immediately adjoin the parts at which the forces are applied, and the symbol B for the remainder. The conditions of equilibrium requires merely that the resultant action transmitted by B shall have a definite

* *Cf.* § 4 and footnote.

† *Cf.* A. E. H. Love, 'Mathematical Theory of Elasticity', § 89.

value; but the conditions of continuity (or of compatability of strains) will not be satisfied unless the total increase of the strain-energy stored in A and B has its minimum value. Evidently, then, the equilibrium configuration may be regarded as in the nature of a compromise between the requirements of A and B. To reduce to a minimum the strain-energy stored in A, the reactions between A and B would require to distribute themselves in a manner which will depend upon the distribution of the forces applied to A; the requirements of B, on the other hand, will not vary, since there must be some definite distribution of stresses, in an otherwise unloaded body, which will entail the minimum storage of strain-energy in transmitting a given resultant action.

Thus, in the process of adjustment which results in the actual distribution corresponding to equilibrium, we may picture a contest between the unloaded portions *, which strive always after "standardization", and the loaded portions, which demand a particular solution for every specified distribution of the forces acting on them. As we pass from the regions of application of load through successive sections of the unloaded portion B, there will be a steady tendency for the claims of standardization to prevail. The theorem stated in italics is an immediate deduction, and we may work back from this, by the principle of superposition, to St. Venant's Principle.

One final remark may be made in this connexion. If our long girder or fuselage is of uniform construction, the requirements of every elementary slice of the unloaded portion, as regards the distribution of stress over its faces which will make the stored strain-energy a minimum, must be the same. We can thus see a physical reason for the predominance of the solutions which St. Venant obtained on the assumption that the stresses are transmitted, in a uniform prism, without change of type: for these must be the "least-energy" solutions for a uniform prism with free surfaces, and as such will be the standard systems to which all particular stress-systems must tend, as we pass from the regions in which the external forces are applied.

* *I. e.*, portions of which the *external* surfaces are free from load.

Part II.

Notation, etc.

9. The remainder of this paper will be concerned with exact analytical proofs of relations which we have previously inferred from physical considerations.

As stated already (§ 6), it is essential that Hooke's law be an expression of the load-extension relation for each member as a whole. We shall find that our equations can be simplified by writing them in terms of "tension-coefficients"*, which we denote by the symbol T: the tension-coefficient for any member is defined as the total tension in that member, divided by its length. In the notation of equation (7), we have then

$$\mathrm{P} = \mathrm{T}l = \Omega l^3 e, \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

where Ω is a constant which conveniently defines the elasticity of the member in question †; and if v denotes the strain-energy stored in any member when it is extended by a tension of magnitude P, we have

$$\left.\begin{aligned} v &= \tfrac{1}{2}\mathrm{P}le &&= \tfrac{1}{2}\mathrm{T}l^2 e, \\ &= \tfrac{1}{2}\frac{\mathrm{P}^2}{l^2\Omega} &&= \tfrac{1}{2}\frac{\mathrm{T}^2}{\Omega}. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (13)$$

The total strain-energy stored in a framework may thus be expressed in the form

$$\left.\begin{aligned} \mathrm{V} &= \Sigma[v], \\ &= \tfrac{1}{2}\Sigma\left[\frac{\mathrm{P}^2}{l^2\Omega}\right] = \tfrac{1}{2}\Sigma\left[\frac{\mathrm{T}^2}{\Omega}\right], \end{aligned}\right\} \quad . \quad . \quad . \quad (14)$$

the summation extending to every member.

In the notation of "tension-coefficients", the equations

* *Cf.* the author's paper on "Primary Stress-Determination in Space Frames", 'Engineering', Feb. 6th, 1920.

† *Cf.* R. and M. 737 of the Aeronautical Research Committee, § 4. In a member of uniform cross-section A, and composed of material of which the Young's Modulus is E, we have

$$\Omega = \frac{\mathrm{EA}}{l^3};$$

the modifications required for taking account of varying cross-section or material are obvious.

of equilibrium for any joint A of the framework take the form *

$$\left.\begin{aligned} X_A &= T_{AB}(x_A - x_B) + T_{AC}(x_A - x_C) + \ldots, \\ Y_A &= T_{AB}(y_A - y_B) + T_{AC}(y_A - y_C) + \ldots, \\ Z_A &= T_{AB}(z_A - z_B) + T_{AC}(z_A - z_C) + \ldots, \end{aligned}\right\} \quad . \quad (15)$$

where X_A, Y_A, Z_A are the components, in the directions of x, y and z respectively, of the external force acting at A, and T_{AB}, T_{AC}, ... etc. are the tension-coefficients for the members, AB, AC, ... etc., which connect the joint A to other joints of the framework.

Let u_A, v_A, w_A denote the component displacements of the joint A, in the directions of x, y and z respectively, from its position in the "zero configuration",—*i. e.*, in the simple and unstrained framework formed by the "essential members" only. Then the fractional increase in the distance between any two joints A and K is given by †

$$\frac{(x_A - x_K)(u_A - u_K) + (y_A - y_K)(v_A - v_K) + (z_A - z_K)(w_A - w_K)}{l_{AK}^2}.$$

Suppose now that the joints A and K are joined by one of the members classified as redundant, and that before insertion this member is "short" by the quantity λ_{AK} of the distance between A and K in the zero configuration. Then, when this member is in place, and the joints A and K have undergone displacements u_A, v_A, w_A, u_K, v_K, w_K, its fractional extension will be given by the expression

$$e_{AK} = \frac{\lambda_{AK}}{l_{AK}} + \frac{(x_A - x_K)(u_A - u_K) + (y_A - y_K)(v_A - v_K) + (z_A - z_K)(w_A - w_K)}{l_{AK}^2}, \quad . \quad . \quad . \quad (16)$$

from which the tension-coefficient for AK may be calculated by means of (12), and inserted in the equations (15) of equilibrium. The expression (16) is clearly applicable to "essential" as well as to "redundant" members, the quantity λ_{AK} being zero for a member of the former type.

* *Cf.* 'Engineering', *loc. cit.*, equations (2)-(4).
† *Cf.* R. and M. 737, equation (2).

The extensions, as usual, are assumed to be so small that their products may be neglected. It is then a matter of indifference whether l_{AK}, in (16), represents the strained or unstrained length of the member AK, or the distance between the joints A and K in the zero configuration.

Proof of the Principle of Superposition (§6).

10. In a framework of any specified dimensions, let u, v, w, ... denote the displacements of the joints from the zero configuration, after equilibrium has been established under a load-system defined by X, Y, Z, ... etc., when the redundant members are initially short by the amounts λ_{AK}, ... etc. And let u', v', w', ... denote the corresponding displacements for a load-system defined by X', Y', Z', ..., when the redundant members are initially short by the amounts λ'_{AK}, ... etc. Then it can be seen at once, from the linearity of the equations (12), (15) and (16), that $(u+u')$, $(v+v')$, $(w+w')$, ... will constitute a solution for the displacements corresponding to a load-system defined by $(X+X')$, $(Y+Y')$, $(Z+Z')$, ... etc., when the redundant members are initially short by the amounts $(\lambda_{AK}+\lambda'_{AK})$, ... etc.

It follows (if we make λ'_{AK}, ... etc. severally zero) that the effect of applying any definite load-system X'_A, Y'_A, Z'_A, ... etc., or (if we make X'_A, Y'_A, Z'_A, ... etc. severally zero) the effect of relaxing our "redundant" members by any definite amounts, is to produce the same increments in the displacements and tensions, whatever be the load-system and self-straining existent initially. This is the Principle of Superposition.

A Reciprocal Theorem.

11. Let T_{AK}, ... etc. denote the tension-coefficients for the members of a given framework, when the initial self-straining is defined by λ_{AK}, ... etc. and the external loads are X_A, Y_A, Z_A, ... etc. And let T'_{AK}, ... etc. denote the tension-coefficients for the same framework, when the initial self-straining is defined by λ'_{AK}, ... etc. and the external loads are X'_A, Y'_A, Z'_A, ... etc. We shall derive alternative expressions for the quantity

$$\Sigma\left[\frac{T.T'}{\Omega}\right], \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

in which the summation extends to every member of the framework.

Let the displacements from the "zero configuration", under the first conditions, be denoted by u_A, v_A, w_A, ... etc. Then we have, from (12) and (16),

$$\frac{T_{AK}}{\Omega_{AK}} = l_{AK}\cdot\lambda_{AK} + \{(x_A-x_K)(u_A-u_K)+(y_A-y_K)(v_A-v_K) + (z_A-z_K)(w_A-w_K)\}. \quad . \quad . \quad . \quad (18)$$

Making this substitution in (17), we have, corresponding to the first term on the right of (18), the quantity

$$\Sigma[l\lambda T'], \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

in which the summation extends to every member. The summation corresponding to the second term on the right of (18) may be arranged in the alternative form

$$\Sigma\left[\begin{array}{l} u_A\{(x_A-x_B)T'_{AB}+\ldots+(x_A-x_K)T'_{AK}+\ldots\} \\ +v_A\{(y_A-y_B)T'_{AB}+\ldots+(y_A-y_K)T'_{AK}+\ldots\} \\ +w_A\{(z_A-z_B)T'_{AB}+\ldots+(z_A-z_K)T'_{AK}+\ldots\} \end{array}\right], \quad . \quad . \quad . \quad (20)$$

where, in the series enclosed by twisted brackets, terms are included for every member which connects the joint A with another joint. The summation in (20) extends to every joint.

But the equations of equilibrium, which are of the type (15), show that (20) may be written in the form

$$\Sigma[uX'+vY'+wZ'],$$

the summation extending to every joint. Hence, and by (19), we have

$$\Sigma_m\left[\frac{T\,.\,T'}{\Omega}\right] = \Sigma_m[l\lambda T'] + \Sigma_j[uX'+vY'+wZ'] = \Sigma_m[l\lambda' T] + \Sigma_j[u'X+v'Y+w'Z], \quad . \quad (21)$$

by symmetry, where suffixes m and j attached to Σ denote summations with respect to every member, and to every joint, respectively.

These alternative expressions for the same quantity give

us a reciprocal relation between any two configurations which can be maintained in a given framework. As a particular example, we may suppose that the self-straining in both configurations is zero. Then $u, v, w, \ldots$ etc. represent the additional displacements of the joints which result from an application of the load-system X, Y, Z, ..., and u', v', w' represent the additional displacements which result from an application of the load-system X', Y', Z', ..., etc. Equations (21) now give the relation

$$\Sigma_m\left[\frac{\mathrm{T}.\mathrm{T}'}{\Omega}\right] = \Sigma_j[\mathrm{X}'u + \mathrm{Y}'v + \mathrm{Z}'w] = \Sigma_j[\mathrm{X}u' + \mathrm{Y}v' + \mathrm{Z}w'], \quad . \quad . \quad . \quad (22)$$

which is Lord Rayleigh's reciprocal theorem *.

Castigliano's First Theorem (§ 4).

12. If we identify the two systems considered in (22), we obtain, as an expression for the strain-energy stored by forces X', Y', Z', ... applied to a framework which is not initially self-strained,

$$\mathrm{V} = \tfrac{1}{2}\Sigma_m\left[\frac{\mathrm{T}'^2}{\Omega}\right] = \tfrac{1}{2}\Sigma_j[\mathrm{X}'u' + \mathrm{Y}'v' + \mathrm{Z}'w']. \quad . \quad (23)$$

Let $\delta u', \delta v', \delta w', \ldots$ etc. denote the increments of displacement, and $\delta\mathrm{V}'$ the increment of strain-energy, resulting from increments $\delta\mathrm{X}', \delta\mathrm{Y}', \delta\mathrm{Z}', \ldots$ made in the applied loads. Regarding all these increments as small quantities, and neglecting small quantities of the second order, we have then, from (23),

$$\delta\mathrm{V} = \tfrac{1}{2}\Sigma_j[\delta\mathrm{X}'.u' + \delta\mathrm{Y}'.v' + \delta\mathrm{Z}'.w'] + \tfrac{1}{2}\Sigma_j[\mathrm{X}'.\delta u' + \mathrm{Y}'.\delta v' + \mathrm{Z}'.\delta w']. \quad (24)$$

But if, in (22), we identify the undashed quantities with the increments now considered, we have

$$\Sigma_j[\mathrm{X}'.\delta u' + \mathrm{Y}'.\delta v' + \mathrm{Z}'.\delta w'] = \Sigma_j[\delta\mathrm{X}'.u' + \delta\mathrm{Y}'.v' + \delta\mathrm{Z}'.w'].$$

Making use of this result in (24), and assuming that all the increments in the applied forces, except one ($\delta\mathrm{X}'_\mathrm{A}$ say),

* *Cf.* A. E. H. Love, 'Mathematical Theory of Elasticity', § 121.

are zero, we obtain the relation

$$\delta V = \delta X'_A . u'_A,$$

whence

$$u'_A = \frac{\partial V}{\partial X'_A}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (25)$$

This is Castigliano's first theorem.

Increase in Strain-Energy caused by the Application of a given Load-System (§ 6).

13. Equation (23) of the preceding section gives an expression for the total strain-energy stored by a load-system X′, Y′, Z′, ... applied to a framework which is not initially loaded or self-strained. We can show that this expression will also hold for the *total increase of strain-energy* which results when the same load-system is applied to a framework which is initially self-strained, provided that no external forces act initially. For if, in the equations (21), T, λ, u, v, w represent a configuration of self-straining without external load (so that X, Y, Z, ... vanish severally), and T′, u', v', w' a system due entirely to external loads X′, Y′, Z′, ... etc. (so that λ′, ... vanish severally), we have

$$\Sigma_m\left[\frac{T \cdot T'}{\Omega}\right] = \Sigma_m[l\lambda T'] + \Sigma_j[uX' + vY' + wZ']$$
$$= 0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (26)$$

Suppose then that the load-system here denoted by X′, Y′, Z′, ... is applied to the framework when it is self-strained to an extent which is defined by the undashed quantities T, ..., λ, ... By the principle of superposition, the tension-coefficients have their values increased to (T + T′), ... etc., and the total increase in the strain-energy

$$= V' = \tfrac{1}{2}\Sigma_m\left[\frac{(T+T')^2}{\Omega}\right] - \tfrac{1}{2}\Sigma_m\left[\frac{T^2}{\Omega}\right], \text{ by (13)},$$
$$= \Sigma_m\left[\frac{T \cdot T'}{\Omega}\right] + \tfrac{1}{2}\Sigma_m\left[\frac{T'^2}{\Omega}\right],$$
$$= 0 + \tfrac{1}{2}\Sigma_m\left[\frac{T'^2}{\Omega}\right], \text{ by (26)}.$$

The expression (23) is thus unaffected by initial self-straining.

Effect of a Relaxation of the Self-Straining.

14. We have shown in § 10 that a system of self-straining may be superposed or subtracted. Let us define any such system by undashed quantities, and imagine it to be subtracted from a configuration represented by barred quantities ($\bar{u}$, $\bar{v}$, $\bar{w}$, $\bar{\lambda}$, $\bar{\mathrm{T}}$, $\bar{\mathrm{X}}$, ... etc.). Then the final value of the total strain-energy is

$$\mathrm{V}'' = \tfrac{1}{2}\Sigma_m\left[\frac{(\bar{\mathrm{T}}-\mathrm{T})^2}{\Omega}\right],$$

$$= \tfrac{1}{2}\Sigma_m\left[\frac{\bar{\mathrm{T}}^2}{\Omega}\right] - \Sigma_m\left[\frac{\bar{\mathrm{T}}.\mathrm{T}}{\Omega}\right] + \tfrac{1}{2}\Sigma_m\left[\frac{\mathrm{T}^2}{\Omega}\right]. \quad . \quad (27)$$

Suppose now that the undashed quantities in (21) relate to a system of self-straining only (so that X, Y, Z, ... etc. vanish identically). Then if we identify the dashed quantities with the undashed, we have

$$\Sigma_m\left[\frac{\mathrm{T}^2}{\Omega}\right] = \Sigma_m[l\lambda\mathrm{T}], \quad . \quad . \quad . \quad . \quad . \quad (28)$$

and if we identify them with our barred quantities, we have

$$\Sigma_m\left[\frac{\mathrm{T}.\bar{\mathrm{T}}}{\Omega}\right] = \Sigma_m[l\lambda\bar{\mathrm{T}}] + \Sigma_j[u\bar{\mathrm{X}} + v\bar{\mathrm{Y}} + w\bar{\mathrm{Z}}]. \quad . \quad (29)$$

We may therefore write (27) in the form

$$\mathrm{V}'' = \mathrm{V} - \Sigma_m[l\lambda\bar{\mathrm{T}}] + \tfrac{1}{2}\Sigma_m[l\lambda\mathrm{T}] - \Sigma_j[u\bar{\mathrm{X}} + v\bar{\mathrm{Y}} + w\bar{\mathrm{Z}}], \quad . \quad (30)$$

where V denotes the initial value of the total stored strain-energy. This is the result obtained in equation (10) of § 7, since in the notation of that section

$$\mathrm{P} = l\bar{\mathrm{T}}$$

and

$$\mathrm{P}_0 = l\mathrm{T}.$$

September, 1922.

XVII. *The Molecular Scattering of Light in Liquid Mixtures.*

By C. V. Raman, *M.A., Hon.D.Sc., Palit Professor of Physics in the Calcutta University, and* K. R. Ramanathan, *M.A., University of Madras Research Scholar* *.

1. *Introduction.*

As is well known, there are many substances, *e. g.* phenol and water, which in the liquid state dissolve each other completely and form a single phase above a certain temperature, but below it are only partially soluble and divide into two coexistent phases. Such mixtures exhibit a marked opalescence or turbidity at temperatures slightly higher than that at which they separate into two layers. Smoluchowski † put forward a thermodynamic explanation of this phenomenon as due to the spontaneous local fluctuations of concentration which occur in the mixture and cause optical inhomogeneity. The problem was further handled by Einstein ‡, who made certain simplifying assumptions, namely, that the liquids are incompressible, that their specific volumes are negligibly small in comparison with those of the saturated vapours emitted by them, and that the latter can be treated as ideal gases, and on this basis derived a formula for the opalescence of the mixture in terms of experimentally determinable magnitudes, that is, the variation of its refractive index and of the partial vapour-pressures of its components with alteration of concentration. The simplifications referred to, however, considerably restrict the validity of the investigation which, in fact, can be regarded as strictly applicable only over a small range of temperature above that at which the mixture separates into two phases. For such a restricted range, the correctness of Einstein's formula has been experimentally established by the investigations of Fürth § and of Zernike ‖. Actually, however, by using carefully purified liquids freed from dust by repeated distillation *in vacuo*, the opalescence in mixtures may be studied over any desired range of temperature and not merely in the immediate vicinity of the critical solution point, and very remarkable and interesting changes in the intensity of the scattered

* Communicated by the Authors.
† *Annalen der Physik*, xxv. p. 219 (1908).
‡ *Annalen der Physik*, xxxiii. p. 1295 (1910).
§ *Wiener Sitzungsberichte*, p. 577, 1915.
‖ Thèse, Amsterdam, 1915.

light and the state of its polarization are observed. These phenomena are also noticeable in mixtures of liquids which are completely soluble in each other at the ordinary temperature. It is found that Einstein's approximate investigation is not adequate to explain the observed facts in these cases and that a stricter investigation is called for. It is proposed in this paper to show how by a modification of the procedure adopted by Einstein, a *perfectly rigorous* expression for the light-scattering can be derived in terms of experimentally determinable quantities. A comparison is also made between the results of the modified theory and those observed in experiment. Specially noteworthy are the effects arising from *molecular anisotropy* which, of course, are not contemplated at all in Einstein's theory. These are discussed in detail in the course of the paper.

2. *Thermodynamical Theory: Application of Boltzmann's Principle.*

In order to deal with the problem under consideration, Einstein assumed that the refractive index of the mixture is a function of a single variable k which expresses the concentration of the mixture, that is the ratio of the mass of the second component present to that of the first component. This is not, however, correct. We have in reality to deal with *two* independent variables, namely, the densities of both components, and the local fluctuations of refractive index are determined by variations of both of these quantities. One way of determining the accidental deviations of density in a mixture is by applying the principles of statistical mechanics. This has been attempted by Ornstein*. In order, however, to express the results in terms of physically measurable quantities, it is simpler to use a purely thermodynamical treatment based upon Boltzmann's well-known relation between entropy and probability. In order to make use of this principle, we have to find the change of entropy resulting from small increases Δm_1 and Δm_2 in the masses, m_1 and m_2 of the components actually present in unit volume of the mixture. The probability of such deviations occurring spontaneously and their mean value may then be evaluated. In order to find the change of entropy involved we have to imagine the masses Δm_1 and Δm_2 introduced into unit volume of the mixture isothermally and by a reversible process. The following thermodynamical operation, which is a modification of that used by Einstein, enables this to be done.

* Proc. Roy. Soc. Amsterdam, xv. p. 54 (1912).

Consider unit volume of the mixture enclosed in a cylinder fitted with a frictionless piston by which the hydrostatic pressure can be varied. The base of the cylinder has two openings, which may be either closed at will or fitted with membranes permeable respectively to the vapour of the first and second components only. A second and much larger reservoir similarly fitted with a piston and semi-permeable openings is assumed to be filled with the liquid mixture, this being maintained at the same temperature and at a hydrostatic pressure equal to the total vapour pressure ϕ of the mixture. The following isothermal operations may then be carried out.

Stage 1.

By means of auxiliary pumps applied to the semi-permeable openings, small masses (Δm_1) and (Δm_2) of the two components may be removed in the state of vapour from the reservoir, the volume of which is assumed to be so great that no appreciable alteration in the composition of its contents occurs in consequence of this removal.

If p_1 and p_2 be respectively the partial pressures of the vapours and v_1 and v_2 their specific volumes, the work done in the process is

$$-p_1v_1\Delta m_1-p_2v_2\Delta m_2+[s_1\Delta m_1+s_2\Delta m_2]\phi, \quad . \quad . \quad (1)$$

where s_1 and s_2 represent respectively the shrinkages of the volume of the mixture by removal of unit masses of the two components.

Stage 2.

Taking the masses (Δm_1) and (Δm_2) of the vapours thus separated, force simultaneously a mass (Δm_1) of the first component and a mass ($k\Delta m_1$) of the second component into the cylinder the contents of which are maintained at a constant hydrostatic pressure ϕ, k being the fraction m_2/m_1. It is clear that in this operation the composition of the mixture, and therefore also its vapour-pressure, remain unaltered. The work done in this process is therefore

$$p_1v_1\Delta m_1+p_2v_2k\Delta m_1-[s_1\Delta m_1+s_2k\Delta m_1]\phi. \quad . \quad . \quad (2)$$

Stage 3.

Force the remaining portion ($\Delta m_2-k\Delta m_1$) of the second component into the cylinder, the hydrostatic pressure in it

being maintained at the equilibrium value of its total vapour-pressure at each instant. The work done in this process is

$$v_2(\Delta m_2 - k\Delta m_1)(p_2 + \tfrac{1}{2}\Delta p_2) - s_2(\Delta m_2 - k\Delta m_1)(\phi + \tfrac{1}{2}\Delta\phi), \quad . \quad (3)$$

where (Δp_2) and $(\Delta\phi)$ represent respectively the changes of the partial vapour-pressure and the total vapour-pressure resulting from the operation.

Stage 4.

The volume of the mixture in the smaller cylinder is now greater than at first by $(s_1\Delta m_1 + s_2\Delta m_2)$. The semi-permeable openings are now closed, and the piston enclosing the contents forced in by increasing the hydrostatic pressure till the original volume is restored. The work done in this process is

$$\tfrac{1}{2}\beta(\Delta p)^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where β is the compressibility, of the mixture and (Δp) is the increase of hydrostatic pressure in the process.

Adding up (1), (2), (3), and (4) we find the total work done is

$$\tfrac{1}{2}(v_2\Delta p_2 - s_2\Delta\phi)(\Delta m_2 - k\Delta m_1) + \tfrac{1}{2}\beta(\Delta p)^2. \quad . \quad . \quad (5)$$

The quantity k is a measure of the concentration of the mixture. If it be altered to $k+\Delta k$ as the result of the operations set out above, it is easily seen that $m_1\Delta k = (\Delta m_2 - k\Delta m_1)$.

Hence (5) may be written in the form

$$\tfrac{1}{2}m_1\left(v_2\frac{\partial p_2}{\partial k} - s_2\frac{\partial\phi}{\partial k}\right)(\Delta k)^2 + \tfrac{1}{2}\beta(\Delta p)^2,$$

or, for brevity,

$$\tfrac{1}{2}\alpha(\Delta k)^2 + \tfrac{1}{2}\beta(\Delta p)^2. \quad . \quad . \quad . \quad . \quad . \quad (6)$$

The next step is to find the mean value of $(\Delta k)^2$ and $(\Delta p)^2$ in the deviations which occur spontaneously in a volume-element dv of the mixture. We may for this purpose use Boltzmann's principle in the form in which it has been expressed by Einstein,

$$dW = Ce^{-\frac{NE}{RT}}d\lambda_1 d\lambda_2 \ldots . d\lambda_n, \quad . \quad . \quad . \quad (7)$$

where $\lambda_1, \lambda_2 \ldots . \lambda_n$ are suitably chosen variables which determine the observable condition of the system and dW is the probability that the magnitudes λ_1, λ_2, etc., lie between λ_1 and $\lambda_1 + d\lambda_1$, λ_2 and $\lambda_2 + d\lambda_2$, etc., R, T, and N have the same significance as in kinetic theory. C is a constant and E is the work necessary to bring the system into the state

under consideration from that of thermodynamical equilibrium. If $\lambda_1, \lambda_2 \ldots$. are assumed to be zero in the state of equilibrium and E can be expressed in the form

$$\mathrm{E} = \tfrac{1}{2}a_1\lambda_1^2 + \tfrac{1}{2}a_2\lambda_2^2 + \ldots. + \tfrac{1}{2}a_n\lambda_n^2, \quad . \quad . \quad . \quad (8)$$

then (7) assumes the familiar form of the Gaussian error-law and it may easily be shown by integration that

$$\left.\begin{aligned} \overline{\lambda_1}^2 = \mathrm{RT}/\mathrm{N}a_1,\ \overline{\lambda_2}^2 = \mathrm{RT}/\mathrm{N}a_2, \text{ etc.,} \\ \overline{\lambda_1\lambda_2} = \overline{\lambda_2\lambda_3} \ldots. = 0, \end{aligned}\right\}, \quad . \quad . \quad (9)$$

where the bars indicate the *mean value* of the quantities written under them.

Comparing (6) and (8), we find from (9) that for an elementary volume dv of the mixture

$$\left.\begin{aligned} (\overline{\Delta k})^2 &= \frac{\mathrm{RT}}{\mathrm{N}}\frac{1}{\alpha dv}, \\ (\overline{\Delta p})^2 &= \frac{\mathrm{RT}}{\mathrm{N}}\frac{1}{\beta dv}, \\ \text{and } \overline{\Delta k}\,\overline{\Delta p} &= 0. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Writing $\epsilon = \mu^2$ where μ is the refractive index of the mixture,

$$\Delta\epsilon = \frac{\partial\epsilon}{\partial k}\cdot\Delta k + \frac{\partial\epsilon}{\partial p}\Delta p,$$

and from (10), it follows at once that

$$(\overline{\Delta\epsilon})^2 = \frac{\mathrm{RT}}{\mathrm{N}dv}\left[\left(\frac{\partial\epsilon}{\partial k}\right)^2\Big/\alpha + \left(\frac{\partial\epsilon}{\partial p}\right)^2\Big/\beta\right]. \quad . \quad . \quad (11)$$

3. *Intensity of the Scattered Light: Comparison with Einstein's formula.*

From (11), the intensity of the opalescence radiation can be easily calculated. Using a relation given by the late Lord Rayleigh *, we find that in a direction at right angles to the incident beam (assumed to be unpolarized and of unit intensity), the intensity of the scattered light is

$$\mathrm{I} = \Sigma\frac{\pi^2(dv)^2(\overline{\Delta\epsilon})^2}{2\lambda^4}\Big/r^2,$$

where λ is the wave-length and r is the distance of the volume-element from the point of observation. Substituting

* Scientific Papers, vol. i. p. 526.

for $(\Delta\epsilon)^2$ from (11), and integrating over the whole volume of the fluid, say Ψ,

$$I=\frac{\Psi}{r^2}\cdot\frac{\pi^2 RT}{2N\lambda^4}\left[\left(\frac{\partial\epsilon}{\partial k}\right)^2\Big/ m_1\left(v_2\frac{\partial p_2}{\partial k}-s_2\frac{\partial\phi}{\partial k}\right)+\beta\rho^2\left(\frac{\partial\epsilon}{\partial\rho}\right)^2\right], \quad . \quad . \quad (12)$$

where ρ is the density of the mixture.

The principal difference between Einstein's expression for the light-scattering in liquid mixtures and that given here is that the former does not include the term proportional to the compressibility of the mixture which appears in (12). This term expresses the light-scattering due to the local thermal *fluctuations of density* in the mixture considered as a single substance. Using Lorentz's relation $(\epsilon-1)/(\epsilon+2)=\text{const.}\rho$, the second term in (12) may be developed and written in the form

$$\frac{\pi^2 RT\beta}{18N\lambda^4}(\mu^2-1)^2(\mu^2+2)^2\frac{\Psi}{r^2}. \quad . \quad . \quad . \quad (13)$$

The first term in (12) which represents the light-scattering due to the local *fluctuations of concentration* in the mixture may be simplified if we neglect $s_2\dfrac{\partial\phi}{\partial k}$ in comparison with $v_2\dfrac{\partial p_2}{\partial k}$ and assume, further, that the saturated vapour obeys the gas laws, in which case it may be written as

$$\frac{M_2}{2N\lambda^4}\cdot\frac{\frac{1}{m_1}(\partial\epsilon/\partial k)^2}{\partial\log p_2/\partial k}\cdot\frac{\Psi}{r^2}, \quad . \quad . \quad . \quad . \quad . \quad (14)$$

where M_2 is the molecular weight of the second component in the gaseous phase. (14) is Einstein's approximate result. The simplifications here made are only permissible when the vapour-pressure is not large.

The relative importance of the two effects, viz. the light-scattering due to fluctuations of density and concentration respectively, which contribute to its total observed intensity depends on the circumstances of the particular case. The first effect depends on the refractive index and compressibility of the mixture, and second on the variations of refractive index and of the partial vapour-pressures produced by a change in the composition of the mixture. In the immediate neighbourhood of the critical solution temperature,

the partial vapour-pressures alter very little with the composition of the mixture, and the concentration-scattering becomes very great. On the other hand, at temperatures far removed from the critical solution point or in the cases in which the components are ordinarily completely miscible, the partial vapour-pressures change rapidly with concentration, the concentration scattering becomes less, and if the refractive indices of the components do not differ much, it may even become unimportant in comparison with the density scattering. The relative proportion of the two components in the mixture is also of importance. It is evident that when either of the components forms a vanishingly small proportion of the mixture, the concentration-scattering would disappear and the observed opalescence would be due entirely to the density-scattering in the remaining component. The composition of the liquid would also influence the density-scattering, since the refractive index and compressibility of the mixture are both functions of it.

4. *Effect of Molecular Anisotropy.*

The theory so far discussed rests on a purely thermodynamical basis and leads to the result that the light scattered in a direction transverse to the primary beam should be completely polarized. If we had proceeded on the basis of a molecular theory we should have obtained a precisely identical result, provided the molecules could be treated as isotropic or spherical scattering particles exerting forces on each other. Actually, however, this is far from being the case. Even in the case of gases, the transversely scattered light is known to be imperfectly polarized, and in liquids at ordinary temperatures this imperfectness of polarization of the scattered light is found to be much more strikingly evident. Taking this into account, we may, in considering the case of liquid mixtures, divide the total observed light-scattering in a direction transverse to the primary beam into three parts, which are assumed to be superposed on each other. A part Ω_1 arises from density fluctuations. A part Ω_2 arises from the concentration fluctuations. Both of these consist of polarized light, and their magnitudes are given by the thermodynamic theory already considered. Another part arises from the varying orientations of the anisotropic molecules, and this consists in great measure of unpolarized light. Denoting by r the ratio of intensity of the weak to the strong component of

polarization in the transversely scattered light, we may write

$$r = \frac{2\Omega_4}{\Omega_1+\Omega_2+\Omega_3+2\Omega_4}, \quad . \quad . \quad . \quad . \quad (15)$$

where $\Omega_3+4\Omega_4$ is the total intensity of the orientation-scattering of which a part Ω_3 consists of plane polarized light and the remaining portion, $4\Omega_4$, consists of unpolarized light, and is therefore equally divided between the two perpendicular directions of vibration. For all practical purposes Ω_3* is negligible, and it remains to determine the magnitude of Ω_4.

If we assume that the molecules in a liquid have their *orientations* entirely at random (though their *positions* are *not*), it is readily seen that the unpolarized part of the orientation-scattering would be simply proportional to the number of molecules n_1 and n_2 of each kind present in unit volume, and we may write

$$\Omega_4 = n_1\chi_1 + n_2\chi_2, \quad . \quad . \quad . \quad . \quad . \quad (16)$$

where χ_1, χ_2 are constants characteristic of each kind of molecule. The magnitude of χ_1 and χ_2 may be determined from (15) and (16) when the values of r corresponding to the case in which the liquid contains only one component or the other are known, and the values of r and Ω_4 for any given composition of the mixture may then readily be calculated. The assumption that the orientation-scattering is simply proportional to the number of molecules per unit of volume is shown to be justified by the fact that in a number of cases it enables the value of r to be successfully predicted for a single substance in the state of liquid when the corresponding value of r for the same substance in the state of vapour is known.

It is evident from (15) that when Ω_1 or Ω_2 is very large, r will be very small. In other words, when either the density scattering or the concentration scattering is very large, the transversely scattered light is almost completely polarized. The latter condition arises when the observations are made in the immediate vicinity of the critical solution temperature. The former condition arises at the plait-point temperature of the mixture, *i. e.* at a temperature which

* In the case of gases obeying Boyle's law, only 1/13 of the orientation-scattering consists of polarized light, as has been shown by Born. In the case of liquids under ordinary conditions, it is probably much smaller on account of the partial correlation of phase of the effects of contiguous molecules.

corresponds to the ordinary critical state for a simple fluid. In either of these cases, the orientation-scattering is a negligible part of the whole observed scattering.

5. *Comparison of Theory and Experiment.*

The theoretical results deduced above agree in their general features with experiment. The scattering of light by mixtures of phenol and water of varying concentrations has been studied in the authors' laboratory by Mr. V. S. Tamma for a range of 30° C. above its critical solution temperature (68° C.). The scattering is exceedingly large in the immediate neighbourhood of the critical solution point, and as the temperature is raised its intensity diminishes, but over the whole range the scattering remains much higher than that of the pure liquids taken together or separately. As is indicated by theory, the polarization of the light scattered in the mixture is far more complete than in pure phenol, which shows a large admixture of unpolarized light. Very near the critical solution temperature, the polarization is practically complete. But as the temperature is raised, it becomes less and less perfect.

Experiments have also been made on mixtures of ethyl alcohol and water. As the percentage of alcohol is increased, the scattering increases to a maximum, then comes down to a minimum and again increases. The extra scattering is largest for a mixture containing about 20 per cent. of alcohol. The case of this mixture is particularly complicated, inasmuch as both its compressibility and refractivity are anomalous. At 20° C., the compressibility is a minimum when the percentage by weight of alcohol is nearly 20. The refractivity of the mixture increases as the alcohol-content is increased, and reaches a maximum when there is about 80 per cent. by weight of alcohol, after which it again diminishes.

The following table gives the density-scattering calculated for different compositions in terms of that of water as a unit.

Percentage of alcohol by weight.	Compressibility at $20° \times 10^6$.	Refractive index at 20° C.	Density scattering.
0	49	1·3375	1·00
10	42	1·3436	0·94
20	40	1·3500	0·94
40	47	1·3600	1·20
60	62·5	1·3670	1·68
80	81	1·3693	2·19
100	100	1·3674	2·72

To calculate the concentration-scattering, we require to know the partial vapour-pressures of the components, and curiously enough, the necessary data do not seem to exist for alcohol-water mixtures. Anyhow, the concentration-scattering would vanish at about 80 per cent. by weight of alcohol, since at that point there is no change of refractivity with concentration, and would be very small over the range 60 to 100 per cent. Since the refractivity of alcohol is not very different from that of water, the effect would not be large at any stage. This agrees with the observed facts.

Measurements of the intensity and polarization of the scattered light in binary mixtures of carbon disulphide and ether, and of benzene and normal hexane have been published by W. H. Martin *. By the use of relation (15), we can calculate the orientation-scattering. Partial vapour-pressure data for the first mixture are available †, and from (14) the concentration-scattering can be calculated. But owing to lack of compressibility data, we cannot calculate the density-scattering. Since, however, this forms only a small fraction of the total scattering in the present case, an error in its estimate would not be of much consequence.

In the attached graph are shown, A the concentration-scattering, B the unpolarized scattering, and C the sum of these two for different concentrations. The values of the total experimental scattering are also plotted in curve D. It will be seen that the curve of orientation-scattering is nearly a straight line, showing, in agreement with (16), that it is proportional to the number of the two kinds of molecules per unit volume. The concentration-scattering was calculated by using (14).

Now

$$\frac{\partial \log p_2}{\partial k} = \frac{1}{p_2}\frac{\partial p_2}{\partial k} = -\frac{x^2}{p_2}\frac{\partial p_2}{\partial x},$$

where x is the mass of the first component (in this case CS_2) in unit mass of the mixture. The values of $\frac{\partial p_2}{\partial x}$ were got from the partial vapour-pressure curve of ether. $\frac{\partial \epsilon}{\partial x}$ was calculated from the formula for the refractive index of mixtures

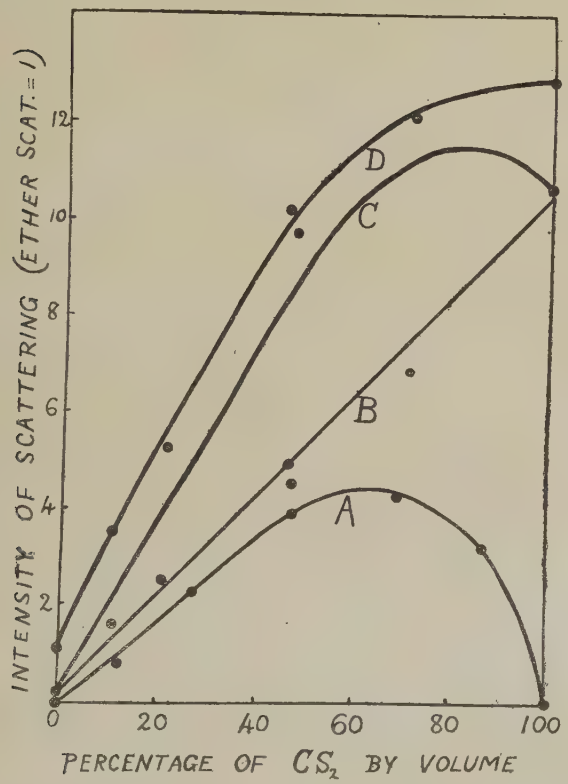
$$\frac{\epsilon-1}{(\epsilon+2)\rho} = \frac{\epsilon_1-1}{(\epsilon_1+2)\rho_1}x + \frac{\epsilon_2-1}{(\epsilon_2+2)\rho_2}(1-x),$$

* Journal of Physical Chemistry, Jan. 1922.

† Guglielmo, *Acc. dei Lincei* (5) 1, i. pp. 242 & 294 (1892).

where ϵ_1, ρ_1, and ϵ_2, ρ_2 refer to the two components of the mixture.

It will be noticed that the curve C closely follows the experimental intensity-curve. The difference, no doubt, represents the density-scattering, which, however, cannot be calculated theoretically (except separately for the two pure liquids) as the compressibility of the mixture is not known.



For pure ether and pure CS_2 the density-scatterings are 0·85 and 2·80, so that if these are added to the orientation-scatterings, the total scattering fits into Martin's experimental curves. To make an adequate test of the theory, we require to know both the compressibilities and the partial vapour-pressures of pairs of liquids. Data regarding partial

vapour-pressures are available for a number of pairs of liquids, and the measurement of compressibilities of mixtures is being undertaken in this laboratory in order to make a complete test of the theory of light-scattering in mixtures.

Summary and Conclusion.

In this paper, the approximate investigation of the scattering of light in liquid mixtures given by Einstein is revised and a perfectly rigorous formula is developed, in which account is taken of the compressibility of the mixture. It is shown that the thermodynamical investigation leads to the result that the light-scattering arises in two distinct ways : (1) due to spontaneous local fluctuations in the composition of the mixture, and (2) due to local fluctuations of its density. Einstein's approximate investigation ignores the second effect, which becomes very important when the observations are made at temperatures not close to the critical solution point or in the case of liquids which are completely miscible at ordinary temperatures.

(*a*) The scattering power is expressed in terms of the compressibility of the mixture, the variations of its refractive index, and the partial vapour-pressures of its constituents with the composition.

(*b*) A discussion is given of the effects of molecular anisotropy which gives rise to an additional "orientation-scattering." It is shown that the polarization of the transversely scattered light is imperfect as the result of molecular anisotropy, but it is considerably less so in the mixture than in the pure liquids, and, further, it becomes more and more perfect as the critical solution temperature is approached.

(*c*) A comparison of the results with the available experimental data shows a general agreement.

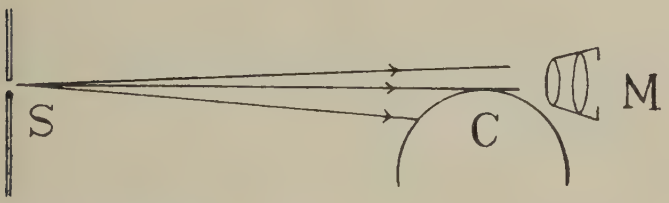
(*d*) Very similar methods may also be applied in the case of solutions of solids in liquids, and in the case of ternary mixtures.

Calcutta,
27th July, 1922.

XVIII. *A Qualitative Study of Inflected Waves.* *By* R. N. Ghosh, *M.Sc.**

IN a recent paper† the author has shown that Airy's theory of inflected wave-fronts and the supernumerary "bows" can be illustrated by ripples. The present note describes a simple optical experiment which qualitatively demonstrates the existence of inflected and cusped wave-fronts.

Fig. 1.



S is the slit which is the source (a Nernst single-filament lamp works well) from which light diverges and falls upon a silvered cylinder of glass C; M is a short-focus microscope objective ($\frac{3}{4}$ in. focal length) so placed that its focal plane is well in advance of the edge of the cylinder grazed by the incident rays. The rays after reflexion at the surface of the cylinder form a virtual caustic within the cylinder, that is the caustic becomes the virtual source of light. Hence as the rays pass through the objective M, they converge or diverge according as they lie away from or within the focal plane of M. After passage through the objective a part of the wave is converging while another is diverging, that is, we have an inflected wave-front just behind the objective. As the inflected wave travels onward, the curvature of the diverging part goes on increasing, while the other half comes to focus first and then diverges, thus folding upon the former half, and giving rise to cusped wave-fronts in the region far behind the objective M. The folding of one part over the other gives rise to interference phenomena. The principal maximum lies in the direction of the cusp where the illumination is maximum. Then we have a series of maxima and

* Communicated by the Author.
† Phil. Mag. vol. xxxv. p. 79 (1918).
‡ Proc. of Indian Assoc. vol. vi. parts iii. & iv. p. 155.

minima of decreasing width and intensity, exactly of the same type as we get in the case of an emerging wave in the rain-drop. In fact, the intensity at any point is given by Airy's integral

$$I=4\left[\int_0^\infty \cos\frac{\pi}{2}(\omega^3-m\omega)\,d\omega\right]^2,$$

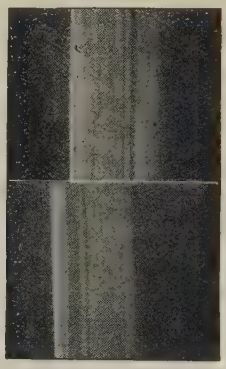
where ω and m have their usual meanings.

The existence of the inflected wave-front can be easily observed by moving a thin vertical wire (·3 mm. diameter) along the wave-front just behind the microscope objective M, and observing the shadow in the region where the cusped wave-fronts are formed. As the wire is moved along the wave-front the shadow first approaches the principal maximum, and retraces its path after reaching the place of maximum illumination.

Since the shadow is in the direction of the wave normal, hence as the wire approaches the point of inflexion of the wave, the shadow moves towards the normal at the point of inflexion

Fig. 2.

Fig. 3.



(which is the direction of the principal maximum), and actually coincides with it at this point; as the wire crosses the point of inflexion the shadow moves away from the normal at the point of inflexion. This is also clear from the formation of the cusped wave-fronts, where one part of the wave on one side of the point of inflexion is folded over the other part. Hence as the wire is moved along the wave, the shadow moves along one part of the cusp, reaches it, and retraces its path along the second part of the cusped wave-front. Fig. 2 shows the shadow before its arrival at the principal maximum,

fig. 3 the same when it has retraced its path after crossing the principal maximum. It will be observed that in fig. 3 the shadow is very faint and broad, the reason being the decrease in the intensity owing to the continued divergence of the part through which the wire passes; while in the other part, it will be remembered, the wave comes to focus first and then diverges, with the result that the illumination in the latter part is much greater than in the other. In figs. 1 and 2 a few diffraction bands are observed to lie on both the sides of the shadow; these diffraction bands are the well-known Fresnel class ones.

The motion of the wire along the wave-front is always accompanied by a disturbance in the interference pattern. (1) The diffraction effects of Fresnel class are always present, but they are almost negligible when the interferences are observed at high magnification. (2) As the wire moves along, the fringe system practically disappears from the region covered by the shadow, because one part of the wave is cut off, and no interference takes place between the two parts over that region.

Some very interesting effects are observed when the wire is just on the point of inflexion of the wave. In order to observe this, care must be taken to get rid of diffraction effects due to the wire by observing the phenomena at a distance of at least 2 metres. The interposition of the wire at the point of inflexion is accompanied by (1) an increase in the number of the fringes, (2) and a diminution in their intensity, (3) and a shift of the principal maximum. In this position of the wire the shadow is not observed, but the principal maximum is replaced by a large number of faint interference fringes, which increase in their number with increasing width of the wire, but greatly diminish in their intensity. The widths of the other maxima are slightly altered. The effect of the wire upon the fringe system diminishes with their order.

Neglecting diffraction effects due to the wire, the intensity at any point can be easily calculated when a thin wire is interposed at the point of inflexion of the wave.

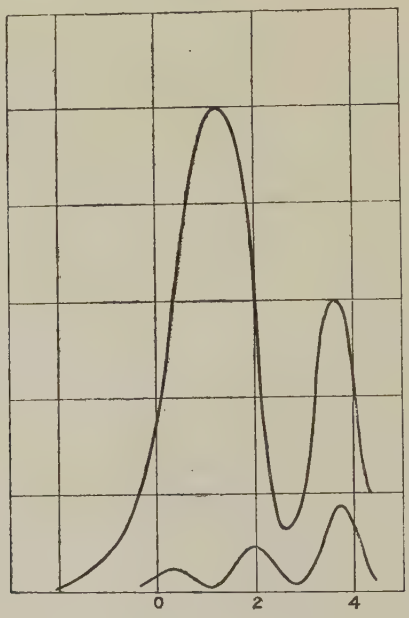
The intensity is then given by

$$I = 4\left[\int_0^\infty \cos\frac{\pi}{2}(\omega^3 - m\omega)d\omega - \int_0^c \cos\frac{\pi}{2}(\omega^3 - m\omega)d\omega\right]^2, \quad (2)$$

where c is the part of the wave cut off by the wire.

In fig. 4 the upper graph represents the intensity when no wire is interposed, while the lower curve represents the intensity when a portion of the inflected wave is cut off at the point of inflexion. The experimental observation is

Fig. 4.



found to agree with the facts brought out by the curve. The increase in the number of fringes, and the diminution of the intensity, and the shift of the principal maximum are clearly shown by the curve. On a future occasion the author hopes to compare the intensities of the interference pattern before and after the interposition of the wire.

The experiments described herein were carried out in the Laboratory of the Muir Central College, Allahabad, and the author begs to record his thanks to the Physics Department, and specially to Principal J. J. Durack for placing the facilities of the laboratory at his disposal and his interest in the work.

Muir Central College,
Allahabad, India.

XIX. *Some Experiments on the Impact of Hydraulic Jets.* *By* Professor A. H. Gibson, *D.Sc.*, *and* F. Heywood, *B.Sc.**

(1) As is well known, the pressure inside the impact orifice of a Pitot tube as commonly used is very closely equivalent to that corresponding to the head of the stream impinging on the orifice. The Pitot tube is, however, almost invariably used to give the velocity at one of a series of points either in an open stream, a jet, or in fluid confined in a pipe, and the area of the orifice is always small compared with that of the stream.

No experiments would appear hitherto to have been carried out to determine what is the maximum size of a Pitot orifice which will correctly register the velocity of a given free jet. Many observers have noted that a free horizontal jet, impinging upon an orifice of its own size in the side of a vessel, is capable of maintaining in that vessel a head of water sensibly equal to the head under which the jet is discharging, which indicates that the area of a Pitot orifice may be at least equal to that of the jet to be measured, but no data are available on the effect of impact on orifices of larger area.

A consideration of the pressures involved in the course of the impact of a free jet on a hemispherical cup would tend to the conclusion that the same pressure head $\left(\frac{v^2}{2g}\right)$ as is recorded in impact on a small orifice, should be attained with an orifice of area considerably greater than that of the jet. For if the cup is of area A and the jet of area a, and if the jet is deflected through 180° on impact, the force exerted on the cup is $\frac{2wav^2}{g}$ lb., and the corresponding pressure per unit of its projected area is $\frac{2wav^2}{Ag}$ lb. or $\frac{2av^2}{Ag}$ ft. of water. So long as A is less than $4a$, this expression is greater than $\frac{v^2}{2g}$, which indicates that the cup is too small to allow of the jet being deflected through 180°. The least size of cup which would theoretically allow the ideal pressure of $\frac{v^2}{2g}$ to be attained with full deviation of the jet, is given by $A=4a$,

* Communicated by the Authors.

and this should therefore be the least size of cup from which the full force of impact may be obtained. With a cup of larger area, the mean pressure intensity would be correspondingly less.

If, then, the phenomena of impact on an open-ended tube full of water is at all analogous to that on a hemispherical cup, it may be expected that so long as the area of this tube is less than about 4 times the area of the jet, the pressure attained in it will be sensibly independent of the area, and approximately equivalent to $\frac{v^2}{2g}$.

(2) No experimental data have hitherto been available as to the limiting area of a hemispherical cup capable of registering the full force of impact, or as to the relationship, if any, between this area and that of the maximum area of an open-ended tube capable of registering the full pressure head of the impinging jet, and the experiments to be described have been carried out in the Whitworth Engineering Laboratories of the University of Manchester with a view of investigating both points.

(3) *Experiments on impact on hemispherical cups.*

The experiments to determine the ratio between the force of impact and the size of cup, were carried out on a series of eleven hemispherical cups having diameters ranging from $\frac{1}{4}$ in. to 1 in. and with jet velocities ranging from 12 f.s. to 60 f.s. These were constructed by pressing hardened steel balls into copper or lead blocks each of which was afterwards machined so as to give a thin sharp-edged cup. This cup was mounted on a beam carefully balanced on knife-edges, and the force of impact of a jet from a parallel bell-mouth orifice vertically below the centre of the cup was measured by being directly balanced by dead weights applied to the beam.

The area and velocity of impact of the jet were obtained from measurements of the diameter of the orifice, the volume discharged, and the difference in level of the jet and the cup.

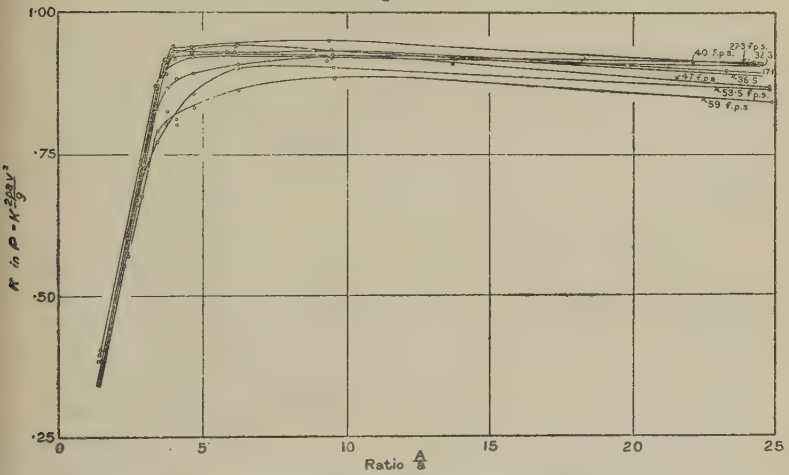
The actual force of impact is always less than the theoretical value $\frac{2wav^2}{g}$ lb., partly because the angle of deviation is never the full 180°. Furthermore if the cup is of large diameter, surface friction reduces the velocity of discharge appreciably, while if of small diameter there is

some interference between the incoming and outgoing jets, and the mean angle of deviation becomes considerably less than 180°. Writing the force of impact as

$$P = k \times \frac{2wav^2}{g} \text{ lb.},$$

fig. 1 shows the relationship between the value of k and the ratio A : a, of the areas of cup and of jet.

Fig. 1.



From this it appears that there is a critical value of A/a, which, for all velocities, lies between 3·6 and 4·0. For smaller values of the ratio the value of k falls off rapidly according to a linear law. As A/a is increased beyond this value, k increases slightly to a maximum, afterwards diminishing, the increase being greater at high than at low velocities. For velocities up to 30 f.s. the maximum value of k is, on the average, only 1 per cent. higher than that obtained when A/a = 4. With higher velocities the value of A/a for maximum efficiency increases with the velocity.

Owing to the flatness of the efficiency curve in the neighbourhood of the maximum point, this point is not well defined, but its value for the jet in question may be taken for all practical purposes as follows :—

Velocity of jet f.s.	20	30	40	50	60
A/a for max. efficiency	8·0	9·0	9·9	10·7	11·5

(4) *Dimensional effect.*

Since viscosity has a comparatively small effect on the phenomena, the principle of dynamic similarity indicates that corresponding efficiencies, and corresponding maximum points, will be attained with larger jets at velocities which are greater than those given above in the ratio of the square roots of corresponding linear dimensions. Thus with a jet 3·25 in. in diameter (sixteen times that of the jet in question) a velocity of 160 ft. per second, equivalent to a head of 400 feet, would correspond to a velocity of 40 f.s. in the small jet, and for this jet diameter and this velocity a ratio $A/a = 9{\cdot}9$ should give maximum efficiency.

Although there are some important differences, the hemispherical cup is to a certain extent analogous to the bucket of a Pelton Wheel, and it is of interest to note that this ratio of bucket to jet area is of the same order as has been found in practice to give best results in the case of the Pelton Wheel.

The principle of dynamic similarity also indicates that since the best ratio of A/a with a given jet increases with the jet velocity, so the best ratio with a given jet velocity should diminish as the size of jet increases. Thus utilizing the foregoing results it appears that whereas at 180 f.s. a 1·8 inch jet should have a bucket ratio of 11·5, a 7·2 inch jet at the same velocity would only require a ratio of 9·0, a result which is in general agreement with Pelton Wheel experience.

(5) *Experiments on impact tubes.*

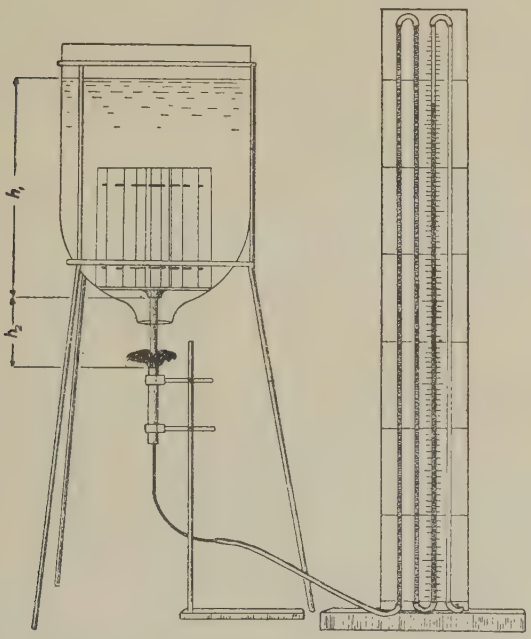
The experiments on the open ended tube consisted in allowing a vertical jet of water from an orifice in the base of a vessel to impinge on an impact tube whose height could be adjusted. A series of orifices were used, these being Borda mouthpieces of diameters 0·3, 0·4, 0·5, 0·575, and 0·7 inch, and a bell-mouthed parallel orifice of 0·5 inch diam. In conjunction with these orifices, impact tubes of the following diameters were used :—0·29, 0·37, 0·42, 0·58, 0·72, 0·84, 0·88, 1·06, and 1·5 inches.

Two series of experiments were carried out, the first with turbulent jets obtained by allowing the Borda mouthpieces to run full, and the second with steady or non-turbulent jets from the convergent mouthpiece.

In each case the impact tube was several diameters in length, was thin-walled, and parallel over its whole length. Its lower end was connected to a water gauge by which the pressure heads were recorded.

The general arrangement is shown in fig. 2. By adjusting either the head above the orifice or the distance of the impact tube from the orifice, the diameter of the jet may readily be adjusted within small limits.

Fig. 2.



Thus if h_1 (fig. 2) denotes the head of water in the containing vessel above the bottom of the orifice,

„ h_2 denotes the distance from the orifice to top of Pitot tube,

„ a_1 denotes the cross-sectional area of the orifice,

„ a_2 denotes the area of the jet at level of Pitot tube,

„ A denotes the area of Pitot tube,

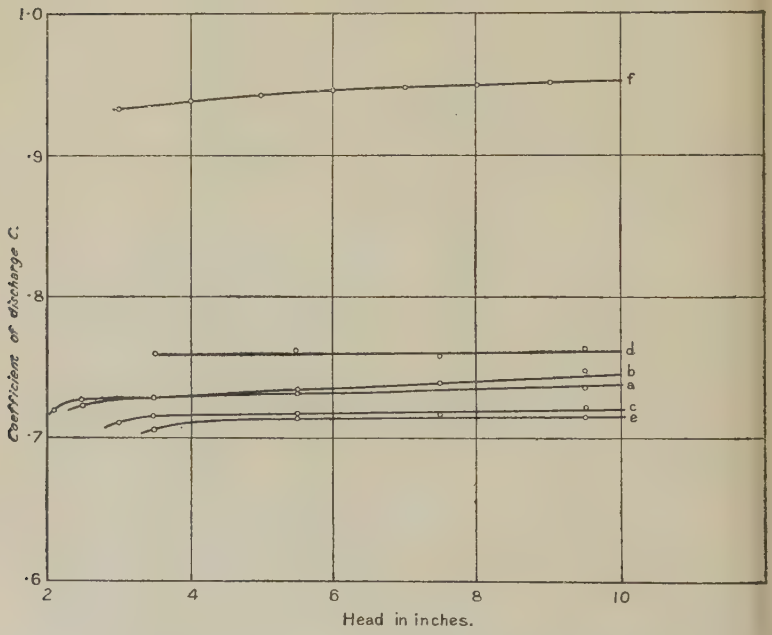
„ C_v denotes the coefficient of discharge for the orifice ;

then
$$a_2 = \frac{C_v a_1 \sqrt{h_1}}{\sqrt{h_2 + C_v^2 h_1}}.$$

The values of C_v for the different orifices were carefully determined, and are shown by the curves of fig. 3.

With the turbulent jets, the recorded pressure is very unsteady, and to obtain consistent readings of the manometer it was necessary to damp out the oscillations by using a long connecting tube of small bore.

Fig. 3.



a. ·3 in. Borda Mouthpiece. *d*. ·575 in. Borda Mouthpiece.
b. ·4 in. „ „ *e*. ·7 in. „ „
c. ·5 in. „ „ *f*. ·5 in. Bell Mouth Orifice.

Expressing the pressure head as $k\frac{v^2}{2g}$, the values of k for the turbulent jets are shown in fig. 4, and for the steady jets in fig. 5. From these curves it appears that so long as the ratio A/a is less than about 3·9 for the steady jets and from 3·4 to 3·7 for the turbulent jets, the value of k is sensibly constant and approximately equal to unity. This is a critical value of the ratio and is practically identical

Fig. 4.

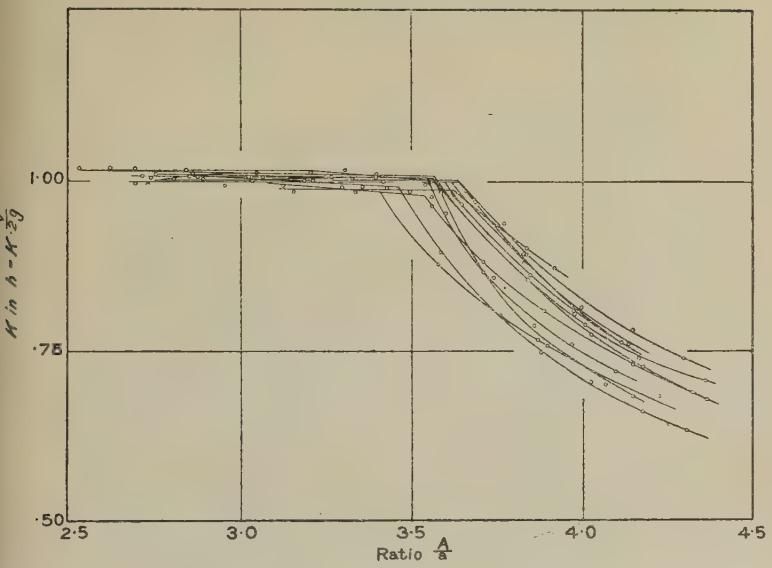
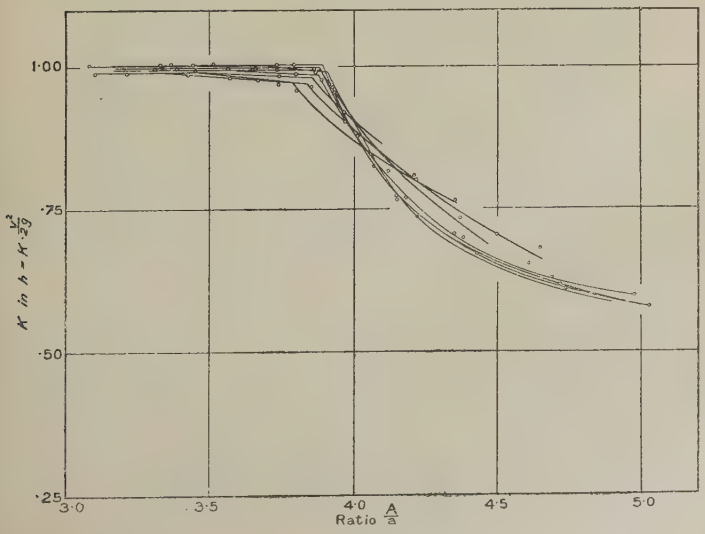
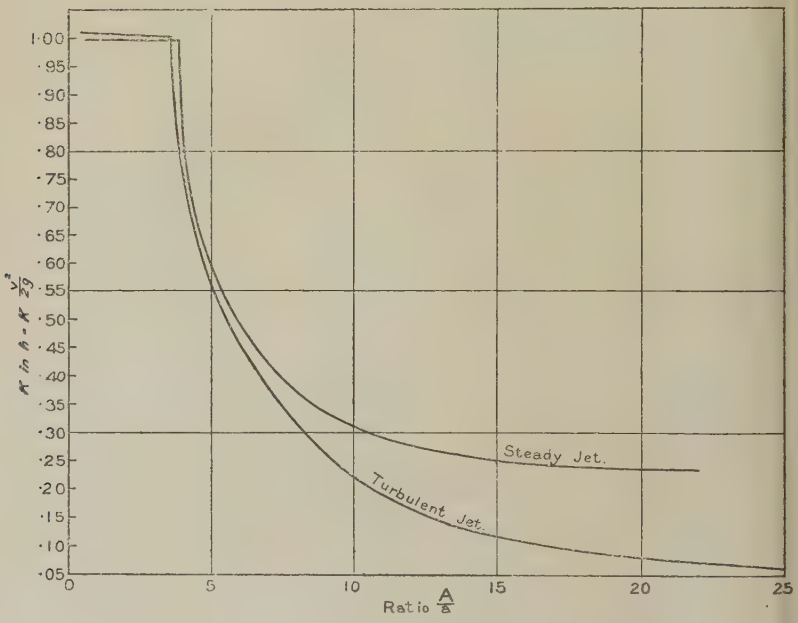

1·00
·75
·50
2·5
3·0
3·5
4·0
4·5
Ratio $\frac{A}{a}$
K in $h - K \cdot \frac{v^2}{2g}$

Fig. 5.


1·00
·75
·50
·25
3·0
3·5
4·0
4·5
5·0
Ratio $\frac{A}{a}$
K in $h - K \cdot \frac{v^2}{2g}$

with the critical ratio found in the case of impact on a hemispherical cup. As A/a is increased beyond this point, k falls rapidly. Its values are shown in the curves of fig. 6.

Fig. 6.



(6) The exact value of the critical ratio is very easily determined in any series of experiments. For values of the ratio smaller than and up to the critical, the water escaping around the periphery of the impact tube takes the form of a continuous clear glassy film, forming a conical surface (fig. 7) whose conical angle diminishes as A/a is increased until it attains a minimum value of 20° for a steady jet and 40° for a turbulent jet*.

The attainment of the critical area-ratio is marked by the breakdown of the film into a series of isolated drops (fig. 8), accompanied by much splashing, while whether with a steady or turbulent jet the pressure in the impact tube becomes very unsteady.

* With the vertex of the cone downwards.

(7) With the object of determining the form of the boundary between the moving and dead water within the impact tube, experiments were carried out with two glass tubes initially filled with aniline dye solution. With a stream-line jet a well-defined boundary between clear and

Fig. 7.

coloured water is soon established (fig. 8), and this is maintained for a considerable time. If l be the distance from the top of the tube to this boundary, and if D be the diameter of the tube and d the diameter of the jet, the value of l/D increases slightly as d increases and also as D increases. If $d/\mathrm{D}=0{\cdot}66$ its value is approximately 3·9 with a 0·4 inch

tube and 4·2 with a 0·56 inch tube, while if $d/D = 0{\cdot}75$ the values are respectively 4·25 and 4·45.

Fig. 8.

With a turbulent jet however, the colour in the tube is instantly dissipated.

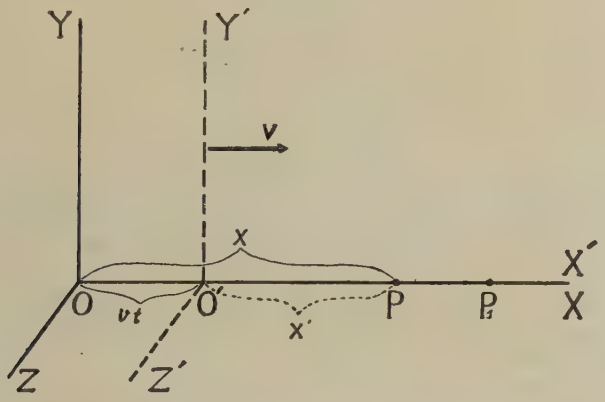
XX. *On a Simple Derivation of the Lorentz Transformations.* *By* H. M. Dadourian, *M.A., Ph.D., Associate Professor of Physics, Trinity College, Hartford, Conn.**

ON account of the important position which the Lorentz Transformations have come to occupy in modern physics it is desirable to make their derivation as direct and simple as possible, to give it a physical rather than a

* Communicated by the Author.

mathematical character, and to bring out clearly the postulates and assumptions which underlie each step. The following derivation of these transformations satisfies these conditions better than any other known to the writer.

Consider the x-coordinate of a point P in space referred to two systems of axes in uniform relative motion and oriented as shown in the following figure. If time is



measured in both systems from the instant the origins of the two systems are coincident, we get

$$x = vt + \mathrm{O'P}.$$

But since space is considered as a homogeneous continuum, the value of O′P obtained with the measuring instruments in the first system is proportional to the value obtained when it is measured with instruments in the second system. Therefore

$$x = vt + kx', \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where the constant of proportionality k is a function of the velocity v.

Similarly $x' = v't' + k'x.$

But two results follow immediately from the first postulate of the restricted theory of relativity :—(*a*) The sense of the mutual velocity is changed without any change in its magnitude when the point of view is changed from one to the other of the two systems, consequently

$$v' = -v.$$

(*b*) The order of points in space is not affected by a change of point of view, consequently k' and k have the same sign. On the other hand, the length of a rod placed in the first system has the same value when observed from the second

system as when the conditions are reversed, consequently k' and k have the same numerical value. Therefore

$$k' = k,$$

and

$$x' = -vt + kx. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

From (1) and (2) we obtain

$$x = \frac{1}{k}(x' + vt'), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

$$t = \frac{1}{k}\left(t' + \frac{1-k^2}{v}x'\right), \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

$$x' = \frac{1}{k}(x - vt), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

$$t' = \frac{1}{k}\left(t - \frac{1-k^2}{v}x\right). \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

In order to determine k we apply the second postulate of the restricted theory to the following special event. A beam of light leaves the common origin of the two systems at the instant $t' = t = 0$ and reaches a point P_1 at the instant the position of P is considered. Then, since the velocity of light is the same in both systems, we have

$$x_1 = ct, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

and

$$x_1' = ct_1'. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

Combining (7) and (8) with (3) and (5) after introducing subscripts into the latter pair to indicate the special character of the event under consideration, we obtain

$$x_1 = \frac{1}{k}\left(1 + \frac{v}{c}\right)x_1', \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

and

$$x_1' = \frac{1}{k}\left(1 - \frac{v}{c}\right)x_1; \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

which give

$$k = \sqrt{1 - \frac{v^2}{c^2}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

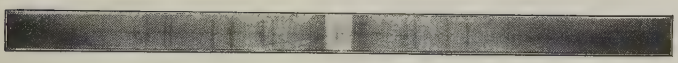
Introducing this value of k in equations (3) to (6) we obtain the famous transformations.

It may be stated parenthetically that the world-points representing the events of the arrival of the beam of light at the points P at the specified instants lie on one of the two light-lines of the xt-plane and satisfy the following conditions :

$$x_1' = \sqrt{\frac{c-v}{c+v}}\,x_1,$$

$$t_1' = \sqrt{\frac{c-v}{c+v}}\,t_1.$$

Fig. 1.



Pd.

Fig. 2.

H_2 absorbed.

Fig. 3.

towards the centre. Hence the following three cases remain possible :—

(i.) The atomic arrangement of palladium atoms in the compound formed is the same as that in the normal state, differing only in their dimensions, and the lines in fig. 2 correspond to this compound.

(ii.) The compound formed is too small in quantity to reveal the lines of sufficient intensity, and the lines in fig. 2 correspond to the palladium remaining in the normal state.

(iii.) The hydrogen is absorbed in the state of solid solution.

The first possibility can be excluded, if we consider the course of the displacement of the lines by changing the percentage of the absorbed hydrogen. A röntgenogram * obtained with a specimen containing a quantity of the hydrogen less than the specimen in fig. 2 showed lines which fell just between the corresponding lines in figs. 1 and 2. If a compound is formed, the transition of the lines must be the gradual disappearance of the lines of palladium in the normal state followed by the appearance of the lines of the compound accompanying this, and not a continuous displacement of the position of lines.

The second case cannot also be possible. In this case the appearance of no new lines can be understood, but we must explain the slight displacement of the lines by the change of the space lattice of the palladium crystal in the specimen. We see from fig. 2 that this change of the lattice is a uniform expansion. This strain necessitates an enormous tension, uniform over all directions. Whatever the volume and form of the compound formed may be, the internal stress in the palladium crystals caused by the production of the compound must be of a very heterogeneous nature, conditioned by the form of the grain boundaries. Hence in this second case the lines must appear as diffused lines, and not as sharp lines displaced, and this possibility is excluded.

On the other hand, if we assume the third case, the results of the experiment can be explained perfectly. Vegard † showed that the solid solution does not reveal any new lines, but a small displacement of lines. Therefore we can without any doubt conclude that the hydrogen is absorbed in the palladium as a solid solution of these two elements.

* Fig. 3 (Pl. II.) shows a few lines of these three röntgenograms. The lines on the middle line are those of fig. 1, the lines on the upper side those of fig. 2, and those at the lower side are the lines of this third röntgenogram.

† L. Vegard, *Zeitschr. für Phys.* v. p. 17 (1921).

The specimen in fig. 2 absorbed nearly 660 times of its volume of the hydrogen gas at room temperature and normal pressure. The molecular percentage of the absorbed quantity was 29 per cent. The quantity of hydrogen was determined by expelling it by heating, after the röntgenogram was obtained. The expansion of the lattice determined by the displacement of the lines was 2·8 per cent, while the expansion of the specimen obtained by the direct measurement of its diameter was 2·9 per cent. This coincidence shows that the expansion of the specimen is due solely to the expansion of the lattice caused by the absorbed hydrogen.

My cordial thanks are due to Prof. K. Honda, under whose kind directions this investigation was carried out.

Iron and Steel Research Institute,
Tôhoku Imperial University, Sendai, Japan.

XXII. *On the Nature and Amount of the Gravitational Deflexion of Light.* *By* Sir Joseph Larmor, *F.R.S.**

1. THE deviation of the rays of light in a gravitational field has proved to be the most searching criterion hitherto available, to discriminate between various ways of forcing gravitation to come within the crucial electrodynamic principle, which assures that velocity of translatory convection of a material system must remain latent in all local physical tests.

The observations at the solar eclipse of 1919 have confirmed on the whole the prediction of the amount of the effect which had been assigned on the basis of the Einstein formulations. But the object of this note is to support a suggestion that the implications of that intricate and beautiful analytical theory have not perhaps even yet been sufficiently explored,—in fact to submit for critical examination an estimate of the deviation of light that is reduced to one half. Though the general theory has been worked over many times in great analytical detail during the last six years, it does not appear that the order of ideas here developed has been closely dealt with.

The guiding considerations are contained in §§ 2, 3; these were originally drafted in anticipation of the recent Solar eclipse, of Sept. 21, which it was hoped would decide, at any rate practically, the questions there submitted. The other sections, some in part memoranda for future constructive

* Communicated by the Author.

application *, have grown up around them from a desire to probe, on all sides, an argument which seems to be in conflict with main dogmatic postulates of the recent relativity theory. Thus on the present view gravitation would not be absorbed into the spatial frame of reference: so the free path of a particle remains a dynamical orbit in the usual sense, and is not merely a straightest path, in the fourfold extension determined by electrodynamics but modified slightly by gravitation. The "principle of equivalence" between a field of gravitation and a field of acceleration would thus seem definitely to fail: indeed its only application was to connect radiation with gravitation.

The argument would therefore affect fundamentally what was the original primary motive of the relativity theory, the effort to dispense with frames of reference other than those directly afforded by the local part of the material system itself. Any modified formulation such as is here submitted must of course be based on the mathematical work of Einstein, and would not in any case impair the value of his resourceful algebraic synthesis, or the general analytical trend of his mode of adapting gravitation into an electrodynamic frame. The intricacies in this subject are however not in mathematical theory, which however complex is yet straightforward; and they are indeed of long standing, going back to the critiques of Leibniz on the *Principia*; in modern form they are concerned with physical interpretations of the same scheme of equations within different analytic frames, including the mode of adaptation from fourfold frames into the practical Newtonian world of astronomy.

This procedure of practical adaptation into Newtonian space and time, even if adequately coherent within its limited range, which fortunately includes the actual world, may be judged timid and unenterprising. The more robust alternative appears to be to postulate that the world of so-called classical physical science is an illusion, an imperfect picture of a reality depending symbolically on a field of fourfold quasi-geometrical algebra, of a presumed cosmos deep below the sensual sources of our concrete physical knowledge, of

* The present paper was completed as it stands (with additional matter) in the middle of last October, except the final part which was drawn up early in November in reply to criticism.

The writer is aware that the weight of expert authority must be against him, and therefore that the chances are not in favour of the correctness of his conclusions: but he still thinks that the subject requires rather urgently to be cleared up, and that early publication of his argument may serve as a first step in that direction.

Compare the destructive criticism of the postulate that orbits are mutual, yet also geodesic curves by J. Le Roux, *Comptes Rendus*, Nov. 6, p. 809, and Dec. 4, p. 1135.

which we may obtain shadowy glimpses, but which is as yet quite beyond any direct grasp or intuition. Atoms of matter would almost disappear as the permanent sources of phenomena: there would be only "world-filaments," in the fourfold, of nearly uniform density, and reality would be merely the pattern of their intersections. The rigour of this procrustean geometric scheme will have to be relaxed, if it is ever to take the place of the flexible formulations of historical physical science.

Yet this fourfold presentation may not be evaded, even though it represents a world entirely foreign to our practical knowledge. For the principle of what may be called the indifference of space, extended after Minkowski and Einstein into the test of isotropic fourfold space-time invariance of the scheme of equations, is imperative. It may be said to take its origin nearly a century ago with Green and MacCullagh, in their general doctrine of isotropic elastic media, real or ideal, developed on Lagrangean principles: in their theoretical constructions the form of the Action function was determined by its local differential invariance, as expressed by the test that the orientation of the local Cartesian spatial lattice-frame of reference could have no influence on its character. The transfer of this principle of spatial invariance into the fourfold extension is the determining criterion in the electrodynamic analysis: and it may perhaps be held to be only accidental that the relevant group of equivalent fourfold frames involves a merging into insignificance of what in the Newtonian world is prominent as local translatory velocity. Where then does this foundation principle of the quasi-geometrical fourfold and its local directional indifference take its origin? If we postulate that the Action is wholly distributed as an Action-density, to be integrated over space and time, that is virtually if we assume that all activity is in a medium of which the particles are merely passive determining singular points, or more precisely, local interfaces of discontinuity which contribute nothing to the Action directly, then the problem of choosing the form of the Action so as to satisfy this condition of invariance becomes a possible and even feasible one, without further limitation. The discussion resolves itself mainly into fixing the nature and extent of the group of transformations for which the invariance subsists. The multiple product $dx\,dy\,dz\,dt$, expressing the differential element of extension involved in the Action integral, must be separately invariant except as regards a factor depending on position. This is obviously secured if δx, δy, δz, δt belong to local coordinates in a fourfold flat extension. But if time were merely an extra dimension of space, the spatial characteristic of locality in

the fourfold would prevent the local time from ranging without limit within a material system limited in space. This is avoided by making time an imaginary instead of a real dimension in the fourfold ; then the attribute of locality persists as regards positions, which can remain adjacent, not as regards time. To establish any quasi-spatial connexion at all between space and time involves an absolute constant c, of the nature of a velocity when interpreted into Newtonian space : the case of Newtonian space and separate time in which alone an idea of velocity is effective appears as the limit as c tends to an infinite value. Thus the fourfold is the natural and only possible further generalization, after the manner of Green, of the principle of minimal Action extended throughout a medium, which is the final unification of Newtonian physics. And the principle of Action maintains its ground as the ultimate unresolvable dynamical unification, as the ideal method for safe and consistent analysis of intricate physical systems.

2. *The relation of gravitation to mass.*—Any approximate analysis must on this order of ideas lead to a field of gravitational potential belonging to each piece of matter, determined say by its mass-inertia m : for the usual velocities of physical astronomy it is to sufficient approximation a radial field, of intensity μ/r for each body, where μ is some property of that piece of matter, perhaps some function of m in default of other suggestions. The equations of motion of another mass m' in this field are taken to be of type

$$m'\ddot{x} = m' \times \text{component of } \mu/r^2.$$

The implication that it is the same m' that occurs on both sides of the equation, and thus disappears from the problem of free orbits (in a specified field) by cancelling, involves the principle of Newton, experimentally tested by him and confirmed with precision by Bessel and finally within 5 parts in 10^8 by Eötvös *, that masses of whatever kinds of matter behave to a pervading gravitational field in a manner determined by the aggregates of their inertias alone, and therefore when free and started alike travel so as to remain together †.

It is a further principle of Newton's formulation in the *Principia* that gravitational attraction between two masses obeys his law of balanced mutual forces : for otherwise as time flows onward the joint motion of the bodies relative to the uniform Newtonian frame would increase without limit.

* The effects on planetary perihelia seem to depend on terms in the equations of rather less than this order of relative magnitude.

† Cf. *Principia*, Leges Motus, Lex iii. cor. 6.

But so far as we have yet gone the attraction of m_1 on m_2 is $\mu_1 m_2/r^2$, while the attraction of m_2 on m_1 is $\mu_2 m_1/r^2$: these cannot balance unless $\mu = \gamma m$, and then the mutual force is the Newtonian $\gamma m_1 m_2/r^2$. The direct evidence (apart from the principle of Action, as *infra*) settling the degree of verified validity of this formula is not terrestrial experiment, but the whole fabric of physical astronomy: the precision of that evidence is in the main far below the Eötvös standard.

In theory this principle requires some formulation purely dynamical: thus the way to get nearer to it is by a dynamical path,—in fact through the principle of minimal Action which, mainly in Hamilton's hands, became the condensed expression of the essentials of general dynamical science.

In the Minkowskian fourfold frame of reference, which is the most effective setting for abstract electrodynamics, in which time runs parallel (and entangled) with spatial coordinates, not of course completely parallel for it is not spatial, the only expression that could be appropriate for the Action A of a free moving particle is

$$A = \int \nu ds, \quad \text{where} \quad ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2,$$

and ν depends on its mass m. The condition of invariance for change of fourfold Cartesian frame restricts A to this form. On expanding in a series, to bring it into line with the forms familiar in ordinary dynamics,

$$A = \int \nu c dt \left(1 - \tfrac{1}{2}\frac{v^2}{c^2} + \dots\right), \quad \text{where } v^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$$

$$= |\nu c t| - \int \left(\frac{1}{2}\frac{\nu}{c} v^2 + \dots\right) dt.$$

When this is interpreted into a Newtonian frame, namely x, y, z and separate t, it should approximate to the Hamiltonian free Action $\int T dt$ where $T = \frac{1}{2} m v^2$. Hence $\nu = -cm$.

If the particle m, instead of being free, moves in a field of force say of potential (energy) U per unit mass, A should approximate for ordinary velocities to $\int (T - W) dt$, where $W = mU$; for that is the form of the Action in the Newtonian frame. This requirement can be secured, with sufficient approximation, very simply, by changing from c to c' in the equation for ds in the Action $A = -\int cm ds$, where

$$c' = c\left(1 + \frac{U}{c^2}\right).$$

This result carries the suggestion that gravitation can be pressed into the Minkowskian electrodynamic fourfold, which

concentrates into intuitive form the Lorentz group of transformations of ordinary space and time, by slight local change of the velocity of light, proportional to the potential of the field at each place. This procedure would insert gravitation, or any other field of force, into the electrodynamic scheme, itself slightly modified thereby : but of course for this very reason, that it does not point to any special law of the field, it is no explanation. Partial or local frames with different values for c' would have to be coherent, that is to merge together into a single continuous universal fourfold frame, just as the local maps of the surface of the Earth have to be made to fit together into a coherent whole. It is the main problem of hypergeometry to determine the integral form for this continuum : a solution could doubtless be effected for any assigned type of field when, on account of the greatness of c, the warping of the frame from flatness is slight.

The next step in this order of tentative development would be the discussion of the Action of a system constituted of two masses, if we could assume it off-hand to be

$$\mathrm{A} = -\int \dot{m}_1 c ds_1 - \int m_2 c ds_2.$$

But this formula is as yet quite without warrant : though it is the necessary foundation of the material tensor that arises in the adjustments of the relativity formulation. When the fields of activity of two physical systems are superposed in the same region, their Actions are not to be simply added without special reason assigned, for they involve quadratic forms just as do the energies.

We have in fact to deal with a physical field of which m_1 is the nucleus, and which in its ramification from m_1 surrounds and influences m_2, or rather influences the field of m_2 which controls m_2 as its own passive nucleus,—and *vice versa*. This involves a single compound field of joint Action-density in the fourfold, determined by m_1 and m_2 as point-poles or nuclei keying up that field, that being a specification of the masses which is sufficient for the present purpose. It is the essential characteristic of this order of ideas, as here developed, that the Action is required to be located in ultimate analysis in the field alone ; thus it is expressed as a fourfold integral, which must be invariant for transformations within the group of frames which all express by means of different coordinates the same essential spatial spread. This requirement as regards form severely restricts, as has been seen, the extent of the group of frames that are permissible, as being equivalent as regards complete free interchange from one to another.

Now in the fourfold extension, everywhere to a first

approximation locally flat, the threefold spatial part has to remain compact or coherent for a wide group of transformations by continuous change, and therefore (Helmholtz, Lie) it has the Euclidean distance-relation ; and the attachment to it of a fourth, though imaginary, dimension *ict*, which is such as must maintain as above invariance of differential fourfold extent, seems to involve that the fourfold also is (differentially) Euclidean. This result appears to be the source and the essential significance of the Minkowskian fourfold representation : it defines, to its utmost possible extent, the group of changes of frame with respect to which an extended distribution of Action, of density purely continuous, could possibly exhibit invariance. Radiation is invariant in the fourfold, for the Action-density of the electrodynamic field is readily shown to be so : and the experimental facts (Michelson and Morley and Miller, Rayleigh and Brace, Trouton and Noble) invite us to force material dynamics and gravitation into this group of equivalent fourfold frames, and so take them out of the different Newtonian group corresponding to c infinite, in which their sciences have formally been developed,—in preference probably to the older reverse procedure of modifying electrodynamics by shrinkage of the sources of its field so as to retain it within a Newtonian frame, which would perhaps in any extensive development be more complex and artificial. Or, from a different point of view : the invariant formulation in terms of pure algebraic quantity is the essential scheme : the element in it which on a Newtonian kinematic interpretation is the local velocity is merged, in an interpretation into the fourfold spread of space-time, with mere local orientation of the inessential Cartesian reference-lattice.

As then has recently been well-known, there is only one spatial differential invariant, necessarily of considerable analytical complexity, that is appropriate thus to serve, though with certain unusual if not anomalous features, as an Action-density in the fourfold expanse, now formulated as of unrestricted Riemann type, varying from place to place ; namely the invariant of local curvature very briefly indicated in outline by Riemann himself for n-fold extension in a formal thesis in Latin, towards the last days of his short life *. The Action-density is thus determined in general type for a gravitational field, if forms of entirely unmanageable complexity are set aside, except as regards a constant

* The ideal would perhaps be a mixed invariant, entangling gravitation with the electric field, instead of an independent spatial form which satisfies experience in the main but leaves gravitation an isolated phenomenon.

multiplier which prescribes scale of magnitude of gravitation. The minimal condition, which may be described as ensuring that the Action is distributed as regards form and density at a uniform level throughout all the extension, then determines restrictions on the general Riemann type of space which may be described as the laws of structure in the gravitational fourfold. It was shown by an approximate analysis, by Einstein and more systematically by Hilbert, that the field thus reduced spatially to minimal Action, is determined by a criterion of Laplacian form, (as was almost inevitable in invariant analysis) so as to be around each mass m of the type specified by the form of ds^2 sketched empirically above (p. 248) as appropriate to gravitation : while the exact analytical solution for the field of a single isolated mass was worked out soon after by Schwarzschild and by Droste.

The Action, after thus being adjusted to minimal form throughout the free extension which is its seat, proves of course to be of very great local density in the region close around one of its determining poles or material aggregations, and thus integrates spatially, when referred to a coordinate frame attached to the pole, to a finite amount which may be described as belonging to that pole ; this spatial integration aggregates the local part of the Action as a time-integral, which it is not necessary to analyze in closer detail except in a discussion of atomic structure ; it constitutes a material term, in addition to the general fourfold distribution of Action which can now ignore the accumulation around the local core. If we denote this separated material part by $-\int mc^2 dt$, then when transferred to an unrestricted frame not attached to the particle, it must change to $-\int mc\,ds$ in order to be of the necessary invariant type. The result for two mass-particles* m_1 and m_2 is

$$\mathrm{A} = -\int m_1 c\,ds_1 - \int m_2 c\,ds_2 + \text{a fourfold integral.}$$

It so happens, very remarkably and even appropriately from our present point of view, that when there is no electric field the density of the latter continuous part, as now adjusted to minimal spatial distribution, is always and everywhere null, so that the whole of the Action is concentrated within and around the pole-structures or particles of matter. The laws of material dynamics are involved in its further minimization, as regards the positions of the particles.

If then the atoms of matter are passive pole-structures,

* In further elucidation, the cognate reduction for the local Action of electrons in motion in an electrodynamic field, in the Maxwellian frame, as effected in 'Æther and Matter' (1900) § 56, may be compared.

singularities of local beknottedness keying up strain of some kind in a surrounding and intervening active medium which is at each moment practically in internal equilibrium, through adjustment by a process analogous to elastic spatial waves, this argument shows that the Actions belonging to them are correctly to be added together, on account of this equilibrium. The only way to evade this indirect conclusion, as expressed in the formula above, would be by the possibility of a mutual, or mixed, invariant term * involving both the masses; it would appear however that any such feasible term must be based on the scalar fourfold product (ds_1, ds_2), which is already recognized as the basis of the electrodynamic Action between two electric charges carried by the particles † and so is excluded from any other interpretation.

3. *The orbital problem for two masses.*—Writing the exact Schwarzschild form for ds^2 (though the original Einstein approximation which only made the measure of time variable would be sufficient for the present purpose) as if each pole were isolated, for the reason above sketched, we have now for the two interacting particles m_1 and m_2 which originate the field,

$$\mathrm{A} = -\int m_1 c \left\{ c^2\left(1-\frac{\mu_2}{r}\right)dt^2 - \left(1-\frac{\mu_2}{r}\right)^{-1} dr^2 - r^2\, d\theta_1^2 - r^2 \sin^2\theta_1\, d\phi_1^2 \right\}^{\frac{1}{2}}$$

$$-\int m_2 c \left\{ c^2\left(1-\frac{\mu_1}{r}\right)dt^2 - \left(1-\frac{\mu_1}{r}\right)^{-1} dr^2 - r^2\, d\theta_2^2 - r^2 \sin^2\theta_2\, d\phi_2^2 \right\}^{\frac{1}{2}},$$

in which r represents r_{12}; and expanding in series so as to change into a form amenable to direct interpretation into a Newtonian frame with (x, y, z) as Euclidean space and separate t,

$$\mathrm{A} = \int dt \left\{ -m_1 c^2 + \tfrac{1}{2}\frac{c^2 m_1 \mu_2}{r} + \tfrac{1}{2} m_1 v_1^2 + \tfrac{1}{2}\frac{m_1 \mu_2}{r}\left(\frac{dr}{dt}\right)^2 + \dots \right\}$$

$$+\int dt \left\{ -m_2 c^2 + \tfrac{1}{2}\frac{c^2 m_2 \mu_1}{r} + \tfrac{1}{2} m_2 v_2^2 + \tfrac{1}{2}\frac{m_2 \mu_1}{r}\left(\frac{dr}{dt}\right)^2 + \dots \right\},$$

in which v_1, v_2 are the velocities of m_1 and m_2 in the frame;

* Any regional distribution of a gravitational momentum, such as makes the compensation in the case of the electrodynamic stress-tensor, at any rate of amount such as would affect the present argument, is wholly excluded by the slowness of the motions of matter.

† Its value is $\int e_1 e_2 (\dot{x}_1\dot{x}_2 + \dot{y}_1\dot{y}_2 + \dot{z}_1\dot{z}_2 - c^2) r_{12}^{-1}\, dt$ where r_{12} is distance in the fourfold, when gravitation is neglected; being symmetrical with respect to both ends, it is in strong contrast to the type of the gravitational Action.

or, retaining only terms of primary importance,

$$A = \ldots + \int dt \left(\tfrac{1}{2} m_1 v_1^2 + \tfrac{1}{2} m_2 v_2^2 + \tfrac{1}{2} c^2 \frac{m_1 \mu_2 + m_2 \mu_1}{r} \right).$$

As the adjusting gravitational pulses in the threefold would travel at the speed of light, it can be assumed as above that the field of each mass is completely adjusted in the space around it at each instant. The attraction, indicated by the potential term, is mutual, as the Newtonian principle requires.

But the Newtonian Action which in the threefold determines in the usual dynamics the mutual elliptic orbits of two masses is

$$A = \int (T - W)\, dt = \int dt \left(\tfrac{1}{2} m_1 v_1^2 + \tfrac{1}{2} m_2 v_2^2 + \frac{\gamma m_1 m_2}{r} + \text{const.} \right)$$

where γ is the constant of gravitation. To identify this with the form at which we have arrived it is necessary to make μ* equal to $\gamma m / c^2$ [not the usual $2\gamma m / c^2$].

Indeed now we are in a position to appeal from the previous less approximate evidence of general physical astronomy to the closer limits afforded by the Eötvös torsion-balance applied to terrestrial centrifugal force, to fortify this result: for it is required by the latter that $\frac{1}{2}(m_1 \mu_2 + m_2 \mu_1)$ shall be proportional to m_1 up to 5 parts in 10^8, which is almost the order of the electrodynamic convectional verifications. But the principle remains an experimental one for which theory as yet assigns no reason.

It is the very indirect way in which gravitation appears, through an arbitrary constant of integration μ within the bracket in A, that is the source of the inadequacy of a formal quasi-spatial tensor theory as regards orbital problems, if the present argument holds.

It may be of interest to observe a parallelism, in this discussion of the gravitational problem of two free bodies, with the premature effort of Helmholtz, in his great memoir (1847) on the Conservation of Energy (as reproduced in Maxwell 'Treatise,' vol. ii. § 544) to derive the interactions of two linear electric currents from the single equation of Energy alone. There also a problem of two variables demanded a wider principle leading to more than one equation, such as would be expressible as minimal Action; and it is the same liability to count a mutual term twice that is thereby avoided.

The question of the frame of reference to which the velocities in this argument are to be referred perhaps demands further attention. In the first place as to the static

* The quantities μ here, namely μ_1 and μ_2 in A, are new entities, identified only ultimately with the μ of § 2: which may be a cause of confusion.

field of an isolated mass: it is the field developed practically instantaneously, referred to a frame practically Newtonian, or Galileian as it is now often named *, which travels along with the mass but without rotation relative to its motion. Alternatively, the reference can be to any equivalent frame, that is, to one convected with uniform translatory velocity relative to the former. It will then be the square of the velocity relative to this new frame that occurs in the approximate expression for the Action; while the potential energy, involving only r, is independent of the frame. The static fields of m_1 and m_2 thus may at first refer to different frames, in which the masses may have any velocities; until by virtue of equivalence within each of two groups of frames we choose a common frame suitable for both moving masses, in which the analysis can be conducted. This comes to asserting that any one of the Newtonian group of frames, differing with respect to uniform translatory velocity, is suitable for expression of the Action of two, or more, masses by simple summation, in sufficient first approximation for practical cases: and the choice that remains open between the members of the group is the expression of the original Newtonian dynamical relativity.

The deviation of a ray of light in the neighbourhood of a large mass m is determined by electrodynamic principles in terms of the local form of the fourfold electrodynamic spatial datum ds^2 as dominated by that mass, which we have found to be

$$ds^2 = c^2\left(1-\frac{\gamma m}{c^2 r}\right)dt^2 - \left(1-\frac{\gamma m}{c^2 r}\right)^{-1} dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta\, d\phi^2.$$

This is different from the form used by Einstein, and by others expounding and developing his theory, in that it involves γ instead of 2γ. The reason is that the constants in this form have been fixed by comparison with the Newtonian Action for the problem of *two* interacting bodies, in which gravitation appears not as a spatial form but as a mutual potential (energy) of their two fields, and must not be counted twice over, once for each.

* The name frames of inertia was proposed long ago by James Thomson, the only brother of Lord Kelvin, who first introduced the idea of equivalent frames of reference: see 'Inertia, Chronometry, Absolute Rotation,' Proc. Roy. Soc. Edin. 1884, or 'Collected Papers,' pp. 379–88. The paper next following, in the same year, discusses a problem whose mere statement is reminiscent of the current gravitational doctrine: given a system of particles in changing mutual configuration, to determine where possible a (moving) frame of constant configuration relative to which each of the particles is in uniform rectilinear motion. In other words, his problem is to determine in what cases the forces of a set of moving particles can be absorbed into a moving frame of reference.

Unless then there is some discrepancy lurking in this development of the dynamical method through analysis of the Action, the deviation of the rays of light should be reduced to half the usually accepted expression. For example, the summary appeal to invariance, in order to fix general forms, which is a main resource of the analytical theory, might conceivably here be misleading. But the systematic deduction from the Action principle seems to provide a security, in this regard, which is absent from a procedure by formal adaptations and adjustments. If the argument proves to be sound, the same reduction to one half would apply to the predicted displacement of periods in solar spectra. But the influence on physical astronomy, which has in fact been used to fix the constants, would remain unchanged *.

An alternative which may conceivably be proposed is to reject dynamics and its relevant Action principle altogether, as transcended by the transfer from Newtonian space and time into fourfold extension ; in fact to work in Einstein's manner with formal adjustable stress-tensors possibly not amenable to that law. But then the theory seems to lose all cohesion ; there would be no reason except empirical convenience, for identifying μ, which enters into the fourfold solution for the field of a point-mass merely as an arbitrary constant of integration and so completely from outside, with either value γm or $2\gamma m$, or indeed for connecting it in any way with m. If then we allow ourselves to contemplate that astronomical observation may not verify the expression, whichever be the appropriate one, for the gravitational deviation of light, what would be the inference ? At the worst, that gravitation would remain, as previously, outside the electrodynamic group of inessential transformations between sets of coordinates belonging to the same one essential spatial fourfold which makes translatory convection undetectable ; and so gravitation would slightly disturb that result.

For at best this beautiful and intricate formulation cannot be held to explain gravitation. To do that it would have to connect the value of γ with other universal constants of atomic structure : whereas, as it stands, γ might quite well (indeed with great increase of simplicity) be zero, when we would have a world without gravitation altogether. The theory merely explains one way by which gravitation, assumed to exist independently, can be modified so as to come under the electrodynamic principle of the insignificance of local convection.

* Its adequacy to explain the outstanding perihelion motion of Mercury has however recently been placed in doubt by a recension of the standard calculations by the astronomer E. Grossmann.

Not the least remarkable feature of the fourfold formulation, as has often been dwelt on in more complicated ways, is the mode of emergence of the Newtonian dynamical scheme of inertia, momentum, and energy of a particle of matter, as the approximate translation into the Newtonian frame of a form of Action $-\int mc\,ds$ in the Minkowskian fourfold which is of a simplicity almost geometric in character. If gravitation should refuse to come in, this striking concatenation of Newtonian dynamics (with modifications for very high velocities) into the simplest possible fourfold quasi-geometrical Action would still persist, in its simpler homogeneous form, in so far as the masses of atoms can be taken to arise from inertia of their local electrodynamic fields. But gravitation would stand outside it as a residual phenomenon, though in essence hardly more unexplained than it is now.

On the other hand, even if there were complete failure of the astronomical tests, all of them very delicate and precarious, as may appear now to be possible, the theory would retain its position as a remarkable provisional consolidation of the field of physical activity, until some more physical explanation including the necessity of gravitation has been formulated, perhaps by its aid.

Summary.—If nothing has been overlooked in this presentation, fundamental consequences seem to be involved. The inclusion of a field of gravitation in the electrodynamic scheme, by making the fourfold extension variable within a principle of dynamical Action, extended throughout it, remains consistent. A stress-tensor, invariant though imperfectly as before, can be duly balanced throughout the field and around the masses; for that result is involved in the Action principle. But the gravitational warping of the fourfold pseudo-space, due to given inertial masses, and all that depends on it, are halved, including the disturbing effects on light. In the Einstein mixed field, electric and gravitating, the balance of the stress-tensor is not disturbed by this halving of the constant of gravitation in it. The orbit of a planet is however no longer a straightest path, determined everywhere locally, in the fourfold extension, as formulated by electrodynamics and its optical rays but modified slightly by gravitation: though it would be so approximately in a different fourfold, but one outside the group for which electrodynamic relativity would subsist. Gravitation therefore cannot be absorbed in the spatial frame of reference. The "principle of equivalence," which in its application has involved only the postulate that light is affected by a field of gravitation in the same manner as by the field of

universal acceleration of matter which it produces, would, along with the implication of a vacuum, not be true. But that postulate seems to have reference to a latent Newtonian frame, for in the fourfold representation the ideas both of gravitation and of acceleration are transcended : having served as a provisional stimulus and a signpost towards a new region to be explored, it might in any case have disappeared from a completed synthesis. However these questions may be settled, the new theories do not (as yet) explain gravitation : the question why it exists at all in the world is as inscrutable still as it was to Newton : but one way has been exhibited along which it can be very slightly and consistently modified so as to bring it into correlation with electrodynamic theory. The same procedure of conscious reversion to practical Newtonian frames of reference is used to explore cognate problems, such as the question of what precisely can be meant by the assertion that the energy of translatory motion of a body whose mass is intrinsic adds to its inertia, and whether such energy gravitates. And other fundamental questions are briefly considered from the same objective point of view.

[Pressure of time in printing prevents more than an abstract of the additional part referred to in footnote, p. 244. The proposition that orbits could still, on Action theory, be determined separately by formulæ of type $-\delta(m_r c ds_r)=0$ in a fourfold extension involving 2γ, while the extension for electrodynamics and optics is a different one involving γ, is general and not confined to the problem of two bodies : but it is otiose, because the object of the theory is to consolidate these two domains. In the electrodynamic fourfold the orbit of a material particle is held to be determined directly by the positions of the other masses in the field, just as the orbit of an ion is already admitted to be determined by the other charged bodies that are present. The approximate expression on p. 252 for the Action, of fourfold type, in the case of two masses can be transformed readily in the known manner to new coordinates, those of their centre of mass and those of their relative position : in the result these sets of variables are found to be separated, and an expression for the Action of the relative orbit thus appears, involving (as Newton proved) a factor $2\gamma(m_1+m_2)$ instead of $2\gamma m_1$ for the mass m_2, whereas the fourfold extension that is the basis of the Action theory, in which optics also must be evolved, involves γm_1 for the field of m_1. The necessity of an underlying uniform Minkowskian fourfold for final reference after the manner of an æther, is discussed.]